# ON THE INVARIANT SYSTEM OF A PAIR OF CONNEXES*

BY

O. E. GLENN

The theoretical part of this paper treats of a transvectant operation between two general connexes

$$\phi = Aa_{1x}\, a_{2x} \cdots a_{rx}\, \alpha_{1u}\, \alpha_{2u} \cdots \alpha_{su},$$

$$\psi = Bb_{1x}\, b_{2x} \cdots b_{\rho x}\, \beta_{1u}\, \beta_{2u} \cdots \beta_{\sigma u},$$

wherein the variables are ternary and $(u)$ contragredient to $(x)$. This operation, defined by polarization, gives all invariant formations of the simultaneous system of two connexes. Our transvectant, of four indices,

$$\tau = (\phi, \psi)^{i,j}_{k,l},$$

can be interpreted in terms of convolution operations employed by Gordan in his memoirs on the systems of the ternary cubic and quartic forms,† and in Maisano's paper‡ on the concomitants of the first five degrees of the ternary quartic. For doubly homogeneous quantics which are in normal form§ (§ 1, (4)) the *Formenreihe* employed by E. Noether‖ includes an ordered succession of terms of transvectants $\tau$. As compared with the methods of the latter author, who dealt with normal ground-forms and therefore with a canonical form of the problem, our algorithm is non-canonical, and therefore symmetrical. The present method bears a closer resemblance to that of Clebsch and Gordan¶ than to subsequently developed methods.

However, as preliminary to the application of the four types of convolution which were employed by all of the before-mentioned authors, we apply to one of the forms $\phi$, $\psi$ to be combined, a single analytic operation of polarization, involving $i$-fold operative factors corresponding to the first type of

* Presented to the Society, February 27, 1915.

† Gordan, Mathematische Annalen, vol. 1 (1869), p. 90, and vol. 17 (1880), p. 217.

‡ Maisano, Giornali di Battaglini, vol. 19.

§ Gordan, Mathematische Annalen, vol. 5 (1873), p. 104, and *Programmschrift* (Leipsic, 1875).

‖ E. Noether, *Ueber die Bildung des Formensystems der ternären biquadratischen Form*, Journal für Mathematik, vol. 134 (1908).

¶ *Ueber biternäre Formen mit contragredienten Variabeln*, Mathematische Annalen, vol. 1 (1869), p. 359.

convolution, $j$-fold operators corresponding to the second type, and $k$- and $l$-fold operators corresponding respectively to the third and fourth types of convolution. The substitutions (1) made in this polar give $\tau$. This secures an ordering of the convolutions corresponding to a given set of indices $i, j, k, l$, so that theorems analogous to those on the difference between two terms of a binary transvectant are readily stated and proved for $\tau$. The preliminary operation of polarization was not employed by Clebsch and Gordan, although they, and likewise E. Noether (loc. cit.), employed the four types of convolution (under different schemes of ordering) to generate the complete systems which they treated.*

Where it is possible to proceed by applying convolutions between two systems which are finite and complete, absolutely or with respect to a set of symbolical moduli, $\tau$ yields all invariant formations (§ 1). Thus if we are given the finite complete system of each connex $\phi, \psi$, then the system derived by transvection between these two systems is itself finite and complete.

In the second part of the paper we employ $\tau$ in deriving a complete system of simultaneous concomitants of a quadric $p_x^2$ and a bilinear connex $a_x\,\alpha_u$ (§ 2). The problem of making reductions within the finite system generated by transvection in this case is treated from the two standpoints of reduction by identities, and by the syzygies connecting the ground forms. It is the author's thought that the present methods present possibilities for extension towards a theory for the general orders when an adequate asyzygetic theory of ternary forms is available.

We give below a list of memoirs on the various known methods of constructing ternary systems.†

## 1. Transvectant systems

Let $\phi, \psi$ be the two general connexes given above. We define the transvectant $\tau$, of four indices, as follows: Polarize $\phi$ by the following operator,

$$\sum \left(y_1^{(1)}\frac{\partial}{\partial x}\right)^{\epsilon_1}\left(y_2^{(1)}\frac{\partial}{\partial x}\right)^{\epsilon_2}\cdots\left(y_\sigma^{(1)}\frac{\partial}{\partial x}\right)^{\epsilon_\sigma}\left(y_1^{(2)}\frac{\partial}{\partial x}\right)^{\iota_1}\left(y_2^{(2)}\frac{\partial}{\partial x}\right)^{\iota_2}\cdots\left(y_\rho^{(2)}\frac{\partial}{\partial x}\right)^{\iota_\rho}$$

$$\times\left(v_1^{(1)}\frac{\partial}{\partial u}\right)^{\sigma_1}\left(v_2^{(1)}\frac{\partial}{\partial u}\right)^{\sigma_2}\cdots\left(v_\rho^{(1)}\frac{\partial}{\partial u}\right)^{\sigma_\rho}\left(v_1^{(2)}\frac{\partial}{\partial u}\right)^{\nu_1}\left(v_2^{(2)}\frac{\partial}{\partial u}\right)^{\nu_2}\cdots\left(v_\sigma^{(2)}\frac{\partial}{\partial u}\right)^{\nu_\sigma},$$

where

$$\left(y_i^{(j)}\frac{\partial}{\partial x}\right) = y_{1i}^{(j)}\frac{\partial}{\partial x_1} + y_{2i}^{(j)}\frac{\partial}{\partial x_2} + y_{3i}^{(j)}\frac{\partial}{\partial x_3},$$

* The introduction of normal ground-forms enabled E. Noether to reduce the number of processes of convolution to two essential types.

† Baker, Cambridge Philosophical Transactions, vol. 15 (1889). Forsyth, American Journal of Mathematics, vol. 12 (1890). White, American Journal of Mathematics, vol. 14 (1892). Mertens, Wiener Berichte, vol. 95.

etc., and $\epsilon_i, \iota_i, \sigma_i, \nu_i = 0$ or $1$, and

$$\sum\epsilon = i, \quad \sum\iota = j, \quad \sum\sigma = k, \quad \sum\nu = l,$$
$$i + j \leqq r, \quad k + l \leqq s, \quad i + l \leqq \sigma, \quad j + k \leqq \rho.$$

Substitute in the resulting polar $(\phi)$

$$(1) \qquad \begin{aligned} &(a) \quad y_p^{(1)} = \beta_p && (p = 1, \cdots, \sigma), \\ &(b) \quad y_p^{(2)} = (b_p\, u) && (p = 1, \cdots, \rho), \\ &(c) \quad v_p^{(1)} = b_p && (p = 1, \cdots, \rho), \\ &(d) \quad v_p^{(2)} = (\beta_p\, x) && (p = 1, \cdots, \sigma), \end{aligned}$$

and multiply each term of the result by the $b_x, \beta_u$ factors whose symbols are not then involved in it. The resulting concomitant we call the transvectant of $\phi$ and $\psi$ of index $\left(\begin{smallmatrix} i, & j \\ k, & l \end{smallmatrix}\right)$ and abbreviate it as

$$\tau = (\phi, \psi)_{k,\, l}^{i,\, j}.$$

We abbreviate the above polar as

$$(\phi) = \sum \Delta_{y^{(1)}}\, \Delta_{y^{(2)}}\, \Delta_{v^{(1)}}\, \Delta_{v^{(2)}}\, \phi.$$

Let $t_1, t_2$ be any two terms of the double mixed polar $(\phi)$. These can written, respectively,

$$t_1 = \gamma_1\, U_1, \qquad t_2 = \gamma_2\, U_2,$$

where $\gamma_i$ $(i = 1, 2)$ are the products of the $a_x, a_y$ factors of $t_i$ $(i = 1, 2)$, and $U_i$ $(i = 1, 2)$ are the products of the $\alpha_u, \alpha_v$ factors of $t_i$ $(i = 1, 2)$, respectively. Leaving out of account the constant factor $A$, $(\phi)$ is an expression of the type

$$(\phi) = \sum (\gamma_1 + \gamma_2 + \cdots)(U_1 + U_2 + \cdots) = t_1 + t_2 + \cdots,$$

wherein

$$\sum\gamma = \frac{\Delta_{y^{(1)}}\, \Delta_{y^{(2)}}\, \phi}{\alpha_{1u} \cdots \alpha_{su}}, \qquad \sum U = \frac{\Delta_{v^{(1)}}\, \Delta_{v^{(2)}}\, \phi}{a_{1x} \cdots a_{rx}}.$$

The first theorem to be proved depends upon the difference $t_1 - t_2$ of two terms of $(\phi)$. We have

$$t_1 - t_2 = \gamma_1(U_1 - U_2) + U_2(\gamma_1 - \gamma_2),$$

and in this each parenthesis is the difference between two terms of a simple mixed polar. Now two terms of a simple polar are said to be *adjacent* when they differ only in that a pair of factors $a_{hy}\, a_{jx}$ in one is replaced in the other by $a_{hx}\, a_{jy}$. By a known theorem we can add and subtract terms within each

parenthesis so that we will have

$$(2)\qquad \begin{aligned} \gamma_1 - \gamma_2 &= [\gamma_1 - \eta'] + [\eta' - \eta''] + \cdots + [\eta^{(i)} - \gamma_2], \\ U_1 - U_2 &= [U_1 - \xi'] + [\xi' - \xi''] + \cdots + [\xi^{(j)} - U_2], \end{aligned}$$

where each bracket is a difference between two adjacent terms of a simple mixed polar.

*Definition.* The sum of the order of a form in the variables $(x)$ plus its order in the variables $(u)$ is called its *grade.*

*Definition.* The process by which we obtain from $\phi = a_{1x} \cdots a_{rx}\, \alpha_{1u} \cdots \alpha_{su}$ any one of the three covariant types

$$c_1 = (a_1\, a_2\, a_3)\, a_{4x} \cdots a_{rx}\, \alpha_{1u} \cdots \alpha_{su},$$

$$c_2 = (\alpha_1\, \alpha_2\, \alpha_3)\, a_{1x} \cdots a_{rx}\, \alpha_{4u} \cdots \alpha_{su},$$

$$c_3 = a_{1a_1}\, a_{2x} \cdots a_{rx}\, \alpha_{2u} \cdots \alpha_{su},$$

is called convolution.

THEOREM 1. *The difference between any two terms $t_1$, $t_2$ of a transvectant*

$$\tau = (\phi, \psi)^{i,j}_{k,l},$$

*equals a sum of terms each of which is a term of a transvectant $\bar{\tau}$ of two forms $\bar{\phi}$, $\bar{\psi}$, at least one of which is of lower grade than the corresponding one in $\tau$. Moreover, $\bar{\phi}$, $\bar{\psi}$ are obtained from $\phi$, $\psi$ respectively by convolution.*

To prove this theorem observe that the differences involved in (2) are essentially of ten types, and except for factors which may be left out of account in pursuing the argument these are

$$\begin{aligned}
&(A)\quad a_{hy_p^{(1)}}\, a_{kx} - a_{ky_p^{(1)}}\, a_{hx}, && (D)\quad a_{hy_p^{(2)}}\, a_{ky_q^{(2)}} - a_{ky_p^{(2)}}\, a_{hy_q^{(2)}},\\
&(B)\quad a_{hy_p^{(1)}}\, a_{ky_q^{(1)}} - a_{ky_p^{(1)}}\, a_{hy_q^{(1)}}, && (E)\quad a_{hy_p^{(2)}}\, a_{kx} - a_{ky_p^{(2)}}\, a_{hx},\\
&(C)\quad a_{hy_p^{(1)}}\, a_{ky_q^{(2)}} - a_{ky_p^{(1)}}\, a_{hy_q^{(2)}}, &&
\end{aligned}$$

$$\begin{aligned}
&(\mathrm{A})\quad \alpha_{hv_p^{(1)}}\, \alpha_{ku} - \alpha_{kv_p^{(1)}}\, \alpha_{hu}, && (\Delta)\quad \alpha_{hv_p^{(2)}}\, \alpha_{kv_q^{(2)}} - \alpha_{kv_p^{(2)}}\, \alpha_{hv_q^{(2)}},\\
&(\mathrm{B})\quad \alpha_{hv_p^{(1)}}\, \alpha_{kv_q^{(1)}} - \alpha_{kv_p^{(1)}}\, \alpha_{hv_q^{(1)}}, && (\mathrm{E})\quad \alpha_{hv_p^{(2)}}\, \alpha_{ku} - \alpha_{kv_p^{(2)}}\, \alpha_{hu}.\\
&(\Gamma)\quad \alpha_{hv_p^{(1)}}\, \alpha_{kv_q^{(2)}} - \alpha_{kv_p^{(1)}}\, \alpha_{hv_q^{(2)}}, &&
\end{aligned}$$

Substitution from (1) in these differences gives all types of expressions which occur in the difference between two terms of the transvectant $\tau$. The results of the substitutions are given below, $(A)$ going into $(A_1)$, $(B)$ into $(B_1)$, etc.

$$(A_1)\quad a_{h\beta_p}\, a_{kx} - a_{k\beta_p}\, a_{hx} = (a_h\, a_k\, (\beta_p\, x)),$$

$$(B_1)\quad a_{h\beta_p}\, a_{k\beta_q} - a_{k\beta_p}\, a_{h\beta_q} = (a_h\, a_k\, (\beta_p\, \beta_q)),$$

$$(C_1)\quad a_{h\beta_p}(a_k b_q u) - a_{k\beta_p}(a_h b_q u) = (a_h a_k b_q)\beta_{pu} - (a_h a_k u) b_{q\beta_p},$$
$$(D_1)\quad (a_h b_p u)(a_k b_q u) - (a_k b_p u)(a_h b_q u) = (a_h a_k u)(b_p b_q u),$$
$$(E_1)\quad (a_h b_p u) a_{kx} - (a_k b_p u) a_{hx} = (a_h a_k u) b_{px} - (a_h a_k b_p) u_x.$$

The five expressions $(A_1)$, $(B_1)$, $(\Gamma_1)$, $(\Delta_1)$, $(E_1)$ are obtainable from the respective expressions $(A_1)$, $(B_1)$, $(C_1)$, $(D_1)$, $(E_1)$ by the substitutions

$$\begin{pmatrix} a_h & a_k & \beta_p & \beta_q & b_p & b_q & u \\ \alpha_h & \alpha_k & b_p & b_q & \beta_p & \beta_q & x \end{pmatrix}. \tag{3}$$

The right-hand sides of these expressions represent terms of transvectants as follows; the particular convolution from $\phi$ in which the factor $a_{hx}\, a_{kx}$ is replaced by $(a_h a_k u)$ being indicated by $(a_h a_k u)\phi'$, and so forth:

$$(A_1)\quad ((a_h a_k u)\phi', \psi)^{i-1, j}_{k, l+1},$$
$$(B_1)\quad ((a_h a_k u)\phi', (\beta_p \beta_q x)\psi')^{i-2, j}_{k+1, l},$$
$$(C_1)\quad ((a_h^* a_k u)\phi', \psi)^{i-1, j-1}_{k+1, l}, \quad \text{and} \quad ((a_h a_k u)\phi', b_{q\beta_p}\psi')^{i-1, j-1}_{k, l},$$
$$(D_1)\quad ((a_h a_k u)\phi', (b_p b_q u)\psi')^{i, j-2}_{k, l},$$
$$(E_1)\quad ((a_h a_k u)\phi', \psi)^{i, j-1}_{k, l}, \quad \text{and} \quad ((a_h a_k u)\phi', \psi)^{i, j-1}_{k+1, l}\, u_x.$$

Inasmuch as $(A_1), \cdots, (E_1)$ arise from $(A_1), \cdots, (E_1)$ by (3), the principle of symmetry in the theory now immediately shows that the theorem is true in all cases, since $(a_h a_k u)\phi'$ is of lower grade than $\phi$ and $(\alpha_h \alpha_k x)\psi'$ is of lower grade than $\psi$.

An immediate corollary is that any monomial concomitant $\kappa$ equals the transvectant $\tau$ of which it is a term plus terms of transvectants of forms one at least of which is a convolution from the corresponding one in $\tau$ and hence of lower grade, times powers of $u_x$.

Evidently a successive application of this corollary gives for $\kappa$ an expansion in the form

$$\kappa = w_0 + w_1 u_x + w_2 u_x^2 + \cdots. \tag{4}$$

Gordan and also Mertens derived a similar result and showed that the $w_i$ were normal forms, i. e. they satisfy the equation

$$\frac{\partial^2 w}{\partial x_1 \partial u_1} + \frac{\partial^2 w}{\partial x_2 \partial u_2} + \frac{\partial^2 w}{\partial x_3 \partial u_3} = 0.$$

THEOREM 2. *Let* $\phi = A_x^r B_u^s$, $\psi = C_x^\rho D_u^\sigma$, *and let* $\tau$ *be the transvectant*

$$\tau = (\phi, \psi)^{i, j}_{k, l}.$$

*Then if it be possible to separate* $\phi$, $\psi$ *into factors,*

$$\phi = \phi_1 \phi_2 = A_{1x}^{r_1} B_{1u}^{s_1} \cdot A_{2x}^{r_2} B_{2u}^{s_2},$$
$$\psi = \psi_1 \psi_2 = C_{1x}^{\rho_1} D_{1u}^{\sigma_1} \cdot C_{2x}^{\rho_2} D_{2u}^{\sigma_2},$$

*and to partition the index numbers,*

$$i = i_1 + i_2, \qquad j = j_1 + j_2, \qquad k = k_1 + k_2, \qquad l = l_1 + l_2,$$

*in such a way that the transvectants*

$$\tau_1 = (\phi_1, \psi_1)^{i_1, j_1}_{k_1, l_1}, \qquad \tau_2 = (\phi_2, \psi_2)^{i_2, j_2}_{k_2, l_2}$$

*exist as separate nonsymbolical entities, then $\tau$ contains reducible terms, namely, terms of the product $\tau_1 \tau_2$.*

The transvectant $\tau$ is evaluated by polarizing $\phi$ by means of the operator

$$\left(y^{(1)}\frac{\partial}{\partial x}\right)^i \left(y^{(2)}\frac{\partial}{\partial x}\right)^j \left(v^{(1)}\frac{\partial}{\partial u}\right)^k \left(v^{(2)}\frac{\partial}{\partial u}\right)^l,$$

and substituting

$$y^{(1)} = D, \quad y^{(2)} = (Cu), \qquad v^{(1)} = C, \qquad v^{(2)} = (Dx).$$

Thus $\tau$ contains the term

$$P\left(y_1^{(1)}\frac{\partial}{\partial x}\right)^{i_1}\left(y_1^{(2)}\frac{\partial}{\partial x}\right)^{j_1} A_{1x}^{r_1}\left(y_2^{(1)}\frac{\partial}{\partial x}\right)^{i_2}\left(y_2^{(2)}\frac{\partial}{\partial x}\right)^{j_2} A_{2x}^{r_2}$$

$$\times\left(v_1^{(1)}\frac{\partial}{\partial u}\right)^{k_1}\left(v_1^{(2)}\frac{\partial}{\partial u}\right)^{l_1} B_{1u}^{s_1}\left(v_2^{(1)}\frac{\partial}{\partial u}\right)^{k_2}\left(v_2^{(2)}\frac{\partial}{\partial u}\right)^{l_2} B_{2u}^{s_2}$$

$$[y_1^{(1)} = D_1, \qquad y_1^{(2)} = (C_1 u), \qquad y_2^{(1)} = D_2, \qquad y_2^{(2)} = (C_2 u),$$
$$v_1^{(1)} = C_1, \qquad v_1^{(2)} = (D_1 x), \qquad v_2^{(1)} = C_2, \qquad v_2^{(2)} = (D_2 x)].$$

This term is a constant times

$$\tau_1 \tau_2 = (\phi_1, \psi_1)^{i_1, j_1}_{k_1, l_1} (\phi_2, \psi_2)^{i_2, j_2}_{k_2, l_2}.$$

Let $[\phi]$ represent a system of connexes given explicitly by

$$\phi_i = a_{ix}^{\omega_i} \alpha_{iu}^{\mu_i} \qquad (i = 1, 2, \cdots),$$

and let $[\psi]$ represent the system

$$\psi_i = b_{ix}^{\eta_i} \beta_{iu}^{\gamma_i} \qquad (i = 1, 2, \cdots).$$

Then a system $[\tau]$ is said to be the system derived from $[\phi]$, $[\psi]$ by transvection when it includes all terms of all transvectants,

$$\tau = (U, V)^{i, j}_{k, l},$$

where $U$ is a product of powers of forms from the system $[\phi]$ and $V$ a product of powers of forms of the system $[\psi]$.

THEOREM 3. *The number of transvectants of the system $[\tau]$ which contain no reducible terms is finite.**

* Gordan proved the finiteness of simultaneous systems in 1875.

In proof let

$$U = \phi_1^{p_1}\phi_2^{p_2}\cdots\phi_a^{p_a}, \qquad V = \psi_1^{q_1}\psi_2^{q_2}\cdots\psi_b^{q_b}.$$

Then the order of $U$ in the $p = p_1 + \cdots + p_a$ distinct $a$ symbols is

$$p_1\omega_1 + p_2\omega_2 + \cdots + p_a\omega_a,$$

and the order of $U$ in the $p$ distinct $\alpha$ symbols is

$$p_1\mu_1 + p_2\mu_2 + \cdots + p_a\mu_a.$$

The total effect of the transvection $\tau$ upon $U, V$ is to produce a sequence of terms by folding $i\beta$ symbols and $jb$ symbols into convolution combinations with $a$ symbols, and by folding $kb$ symbols and $l\beta$ symbols into convolution combinations with $\alpha$ symbols. In one of these terms $t_1$ of $\tau$ let the number of $a$ symbols not in convolution combination with other symbols be $\rho_1$. Likewise let the number of $\alpha$ symbols of $t_1$ not so combined be $\rho_2$, the number of $b$ symbols not combined be $\rho_3$, and the number of $\beta$ symbols not combined with other symbols be $\rho_4$. Then we have

$$\begin{aligned}
p_1\omega_1 + p_2\omega_2 + \cdots + p_a\omega_a &= \rho_1 + i + j,\\
p_1\mu_1 + p_2\mu_2 + \cdots + p_a\mu_a &= \rho_2 + k + l,\\
q_1\eta_1 + q_2\eta_2 + \cdots + q_b\eta_b &= \rho_3 + j + k,\\
q_1\gamma_1 + q_2\gamma_2 + \cdots + q_b\gamma_b &= \rho_4 + i + l.
\end{aligned} \tag{5}$$

We thus have a system of four linear diophantine equations in the variables $p_1, \cdots, p_a, q_1, \cdots, q_b, \rho_1, \cdots, \rho_4, i, j, k, l$, to be satisfied in positive integers. Each reducible solution of this system corresponds to a transvectant $\tau$ having reducible terms, by the theorem last proved. But the number of irreducible solutions of such a system is finite by a well-known theorem. Hence the theorem is proved.

*Definition.* A system of connexes containing an infinite number of individuals is said to be *finite* when there exists a finite set of the connexes such that every connex of the system is a rational integral function of the members of the set.

*Definition.* A system is *complete* when any form obtained from a product of connexes of the system, by convolution, is expressible rationally and integrally in terms of connexes of the system.

A proof of a theorem now follows readily from the properties of transvectants established above, giving a method of constructing the finite system of a pair of connexes.

THEOREM 4. *If two infinite systems of connexes* $[\phi]$, $[\psi]$ *are each finite and complete then the system* $[\tau]$ *derived therefrom by transvection is finite and complete.*

To prove the finiteness we have only to arrange the transvectants of $[\tau]$ in an ordered sequence arranged according to ascending degrees, those of the same degree being ordered according to ascending grade of the product $UV$, where the typical transvectant of the sequence is

$$\tau = (U, V)^{i,j}_{k,l}.$$

Since the difference between any two terms of $\tau$ is a sum of terms of transvectants,

$$\bar{\tau} = (\bar{U}, \bar{V})^{i',j'}_{k',l'},$$

where $\bar{U}\bar{V}$ is of lower grade than $UV$, a set of connexes in terms of which all others of $[\tau]$ may be rationally expressed may be obtained by selecting terms of the transvectants in this sequence, and precisely one term from each transvectant which contains no reducible term. But the number of transvectants in $[\tau]$ which contain no reducible terms is finite.

To prove completeness let the finite set described above be $K_1, \cdots, K_t$, and let $\bar{P}$ be any monomial form obtained by convolution from

$$P = K_1^{k_1} K_2^{k_2} \cdots K_t^{k_t}.$$

Then $\bar{P}$ is a term of a transvectant

$$(\bar{U}, \bar{V})^{i,j}_{k,l} u_x^s,$$

and in which $\bar{U}$ is obtainable by convolution from $[\phi]$ forms alone, and $\bar{V}$ from $[\psi]$ forms alone. Hence

$$\bar{P} = u_x^s (\bar{U}, \bar{V})^{i,j}_{k,l} + \sum u_x^{s'} (\bar{U}, \bar{V})^{i',j'}_{k',l'}.$$

But $[\phi]$, $[\psi]$ are, by assumption, complete systems, and consequently all transvectants on the right belong to the system $[\tau]$. Hence $[\tau]$ is complete.

## 2. Fundamental system of a conic and a bilinear connex

We let $F = p_x^2 = q_x^2 = \cdots$ be a quadric, $L = (pqu)^2$ its contravariant, $D = (pqr)^2$ its discriminant. The connex $f = a_x \alpha_u = b_x \beta_u = \cdots$ has the fundamental system*

$$f, f_1 = a_x \beta_u b_a, \qquad g = a_x c_x (\beta\gamma x) b_a, \qquad h = \beta_u \gamma_u (acu) b_a,$$

$$i = a_a, \qquad i_1 = a_\beta b_a, \qquad i_2 = a_\beta b_\gamma c_a.$$

The simultaneous system of $F$ and $f$ is the transvectant system given by

$$(F^r L^s, f^t f_1^u g^v h^w)^{i,j}_{k,l},$$

together with the invariants $i$, $i_1$, $i_2$, $D$. There are, in addition to the

* Clebsch and Gordan, Mathematische Annalen, vol. 1 (1869), p. 359.

members of the two individual systems, precisely 92 transvectants of this formula not immediately reducible by theorem (2) of § 1. This is exclusive of the large number which are reducible by the known syzygy

$$gh = \begin{vmatrix} i''f - ii''u_x + i'f_2 & i''u_x & f_2 \\ i''u_x & if - i'u_x + f_2 & f \\ f_2 & f & u_x \end{vmatrix};$$

$$i' = \tfrac{1}{2}(i^2 - i_1), \qquad i'' = \tfrac{1}{6}(i^3 - 3ii_1 + 2i_2), \qquad f_2 = f_1 - if + i'u_x.$$

Of the 92, twenty-three are reducible by the identity

$$(6) \qquad (abu)(\rho\sigma x) = \begin{vmatrix} a_\rho & a_\sigma & a_x \\ b_\rho & b_\sigma & b_x \\ \rho_u & \sigma_u & u_x \end{vmatrix},$$

which is applicable in most cases where $l > 0$. The remaining 69 forms are arranged in the following table according to ascending grade of $UV$, in accordance with the theory.

TABLE

*Grade 4*

$$(F,f)^{10}_{00} = p_a p_x a_x, \qquad (F,f)^{01}_{00} = (pau)\alpha_u p_x,$$
$$(F,f)^{11}_{00} = (pau)p_a, \qquad (F,f_1)^{10}_{00} = p_\beta b_a p_x a_x,$$
$$(F,f_1)^{01}_{00} = b_a(pau)\beta_u p_x, \qquad (F,f_1)^{11}_{00} = p_\beta b_a(pau),$$
$$(L,f)^{00}_{10} = (pqa)(pqu)\alpha_u, \qquad (L,f)^{00}_{01} = (pqu)(\overline{pq}\alpha x)a_x,$$
$$(L,f)^{00}_{11} = (pqa)(\overline{pq}\alpha x), \qquad (L,f_1)^{00}_{10} = b_a(pqa)(pqu)\beta_u,$$
$$(L,f_1)^{00}_{01} = b_a(pqu)(\overline{pq}\beta x)a_x, \qquad (L,f_1)^{00}_{11} = b_a(pqa)(\overline{pq}\beta x).$$

*Grade 5*

$$(F,h)^{10}_{00} = b_a p_\beta (acu)\gamma_u p_x, \qquad (F,h)^{20}_{00} = b_a p_\beta p_\gamma (acu),$$
$$(L,g)^{00}_{10} = b_a(pqa)(pqu)(\beta\gamma x)c_x, \qquad (L,g)^{00}_{20} = b_a(pqa)(pqc)(\beta\gamma x).$$

*Grade 6*

$$(FL,f)^{10}_{10} = p_a(qra)(qru)p_x, \qquad (FL,f_1)^{10}_{10} = b_a p_\beta (qra)(qru)p_x,$$
$$(F,f^2)^{20}_{00} = p_a p_\beta a_x b_x, \qquad (F,f^2)^{02}_{00} = (pau)(pbu)\alpha_u\beta_u,$$
$$(F,f_1^2)^{20}_{00} = b_a p_\beta p_\delta d_\gamma a_x c_x, \qquad (F,f_1^2)^{02}_{00} = b_a d_\gamma (pau)(pcu)\beta_u\delta_u,$$
$$(F,ff_1)^{20}_{00} = b_a p_\beta p_\gamma a_x c_x, \qquad (F,ff_1)^{02}_{00} = b_a(pau)(pcu)\beta_u\gamma_u,$$
$$(L,f^2)^{00}_{20} = (pqa)(pqb)\alpha_u\beta_u, \qquad (L,f^2)^{00}_{02} = (\overline{pq}\alpha x)(\overline{pq}\beta x)a_x b_x,$$
$$(L,f_1^2)^{00}_{20} = b_a d_\gamma (pqa)(pqc)\beta_u\delta_u, \qquad (L,f_1^2)^{00}_{02} = b_a d_\gamma (\overline{pq}\beta x)(\overline{pq}\delta x)a_x c_x,$$
$$(L,ff_1)^{00}_{20} = b_a(pqa)(pqc)\beta_u\gamma_u, \qquad (L,ff_1)^{00}_{02} = b_a(\overline{pq}\beta x)(\overline{pq}\gamma x)a_x c_x.$$

*Grade 7*

$$(F^2, g)^{03}_{00} = b_\alpha (pau)(pcu)(q\overline{\beta\gamma u})\, q_x,$$
$$(F^2, h)^{30}_{00} = b_\alpha p_\beta p_\gamma (qac)\, q_x,$$
$$(L^2, g)^{00}_{30} = b_\alpha (pqa)(pqc)(rs\overline{\beta\gamma})(rsu),$$
$$(L^2, h)^{00}_{03} = b_\alpha (\overline{pq}\beta x)(\overline{pq}\gamma x)(\overline{rs}\overline{ac}x)(rsu),$$
$$(LF, g)^{01}_{20} = b_\alpha (qrc)(qr\overline{\beta\gamma})(pau)\, p_x,$$
$$(FL, h)^{20}_{01} = b_\alpha p_\beta p_\gamma (qru)(\overline{qr}\overline{ac}x).$$

*Grade 8*

$$(FL, f^2)^{20}_{10} = p_\alpha p_\beta (qra)(qru)\, b_x,$$
$$(FL, f^2)^{20}_{20} = p_\alpha p_\beta (qra)(qrb),$$
$$(FL, f^2)^{10}_{20} = p_\alpha (qra)(qrb)\, p_x \beta_u,$$
$$(FL, f_1^2)^{20}_{10} = b_\alpha p_\beta p_\delta d_\gamma (qra)(qru)\, c_x,$$
$$(FL, f_1^2)^{20}_{20} = b_\alpha p_\beta p_\delta d_\gamma (qra)(qrc),$$
$$(FL, f_1^2)^{10}_{20} = b_\alpha p_\beta d_\gamma (qra)(qrc)\, \delta_u p_x,$$
$$(FL, ff_1)^{20}_{10} = b_\alpha p_\beta p_\gamma (qra)(qru)\, c_x,$$
$$(FL, ff_1)^{20}_{20} = b_\alpha p_\beta p_\gamma (qra)(qrc),$$
$$(FL, ff_1)^{10}_{20} = b_\alpha p_\beta (qra)(qrc)\, \gamma_u p_x.$$

*Grade 9*

$$(F^2, fg)^{04}_{00} = b_\alpha (pau)(pcu)(qdu)(q\overline{\beta\gamma u})\, \delta_u,$$
$$(F^2, fh)^{40}_{00} = b_\alpha p_\beta p_\gamma q_\delta (qac)\, d_x,$$
$$(F^2, f_1 g)^{04}_{00} = b_\alpha e_\delta (pau)(pcu)(qdu)(q\overline{\beta\gamma u})\, \epsilon_u,$$
$$(F^2, f_1 h)^{40}_{00} = b_\alpha e_\delta p_\beta p_\gamma q_\epsilon (qac)\, d_x,$$
$$(L^2, fg)^{00}_{40} = b_\alpha (pqa)(pqc)(rsd)(rs\overline{\beta\gamma})\, \delta_u,$$
$$(L^2, fh)^{00}_{04} = b_\alpha (\overline{pq}\beta x)(\overline{pq}\gamma x)(\overline{rs}\delta x)(\overline{rs}\overline{ac}x)\, d_x,$$
$$(L^2, f_1 g)^{00}_{40} = b_\alpha e_\delta (pqa)(pqc)(rsd)(rs\overline{\beta\gamma})\, \epsilon_u,$$
$$(L^2, f_1 h)^{00}_{04} = b_\alpha e_\delta (\overline{pq}\beta x)(\overline{pq}\gamma x)(\overline{rs}\epsilon x)(\overline{rs}\overline{ac}x)\, d_x,$$
$$(FL, fg)^{02}_{20} = b_\alpha (qrd)(qr\overline{\beta\gamma})(pau)(pcu)\, \delta_u,$$
$$(FL, fh)^{20}_{02} = b_\alpha p_\beta p_\gamma (\overline{qr}\delta x)(\overline{qr}\overline{ac}x)\, d_x,$$
$$(FL, f_1 g)^{02}_{20} = b_\alpha e_\delta (qrd)(qr\overline{\beta\gamma})(pau)(pcu)\, \epsilon_u,$$
$$(FL, f_1 h)^{20}_{02} = b_\alpha e_\delta p_\beta p_\gamma (\overline{qr}\epsilon x)(\overline{qr}\overline{ac}x)\, d_x.$$

*Grade 10: No irreducible forms*

*Grade 11*

$$(F^2L, fh)^{40}_{10} = b_\alpha p_\beta p_\gamma q_\delta (qac)(rsd)(rsu),$$
$$(FL^2, fg)^{10}_{40} = b_\alpha (qra)(qrc)(std)(st\overline{\beta\gamma})\, p_\delta p_\epsilon,$$
$$(F^2L, f_1 h)^{40}_{10} = b_\alpha e_\delta p_\beta p_\gamma q_\epsilon (qac)(rsd)(rsu),$$
$$(FL^2, f_1 g)^{10}_{40} = b_\alpha e_\delta (qra)(qrc)(std)(st\overline{\beta\gamma})\, p_\epsilon p_\epsilon.$$

*Grade 12*

$$(F^3, g^2)^{06}_{00} = b_\alpha e_\delta (pau)(pcu)(qdu)(qfu)(r\overline{\beta\gamma u})(r\overline{\epsilon\eta u}),$$
$$(L^3, g^2)^{00}_{60} = b_\alpha e_\delta (opa)(opc)(qrd)(qrf)(\overline{st}\beta\gamma)(\overline{st}\epsilon\eta),$$
$$(F^3, h^2)^{60}_{00} = b_\alpha e_\delta p_\beta p_\gamma q_\epsilon q_\eta (rac)(rdf),$$
$$(L^3, h^2)^{00}_{06} = b_\alpha e_\delta (\overline{op}\beta x)(\overline{op}\gamma x)(\overline{qr}\epsilon x)(\overline{qr}\eta x)(\overline{st}\overline{ac}x)(\overline{st}\overline{df}x),$$
$$(F^2 L, g^2)^{04}_{20} = b_\alpha e_\delta (rs\overline{\beta\gamma})(rs\overline{\epsilon\eta})(\underline{pa}u)(pcu)(qdu)(qfu),$$
$$(FL^2, g^2)^{02}_{40} = b_\alpha e_\delta (qrd)(qrf)(st\overline{\beta\gamma})(st\overline{\epsilon\eta})(pau)(pcu),$$
$$(F^2 L, h^2)^{40}_{02} = b_\alpha e_\delta p_\beta p_\gamma q_\epsilon q_\eta (\overline{rsac}x)(\overline{rsdf}x),$$
$$(FL^2, h^2)^{20}_{04} = b_\alpha e_\delta p_\beta p_\gamma (\overline{qr}\epsilon x)(\overline{qr}\eta x)(\overline{st}\overline{ac}x)(\overline{st}\overline{df}x).$$

In order to illustrate the process of reduction of the transvectants which are reducible by (6) but not by theorem (2), § 1, let us consider

$$\tau = (FL, h)^{10}_{01}.$$

This has the term

$$t = b_\alpha p_\beta (acu)(qru)(\overline{qr}\gamma x) p_x,$$

and

$$t = b_\alpha p_\beta (qru) p_x \begin{vmatrix} (qra) & a_\gamma & a_x \\ (qrc) & c_\gamma & c_x \\ (qru) & \gamma_u & u_x \end{vmatrix}.$$

We treat each term of this $t$ separately. Thus we deduce

$$(qra)(qru) b_\alpha p_\beta p_x c_\gamma u_x = i \cdot u_x \cdot (FL, f_1)^{10}_{10},$$
$$(qrc)(qru)\gamma_u b_\alpha p_\beta p_x a_x = (L, f)^{00}_{10} \cdot (F, f_1)^{10}_{00},$$
$$(qru)^2 b_\alpha a_\gamma p_\beta p_x c_x = L \cdot (F, a_\gamma \beta_u c_x b_\alpha)^{10}_{00}.$$

In the latter case we have*

$$a_\gamma \beta_u b_\alpha c_x = \tfrac{1}{6} u_x (i^3 + 2i_2 - 3ii_1) - \tfrac{1}{2}(i^2 f - i_1 f - 2if_1).$$

Hence $(F, a_\gamma \beta_u c_x b_\alpha)^{10}_{00}$ is reducible in terms of invariants and

$$(F, f)^{10}_{00}, \qquad (F, f_1)^{10}_{00}, \qquad F.$$

Furthermore,

$$(qru)^2 c_\gamma b_\alpha p_\beta p_x a_x = i \cdot L \cdot (F, f_1)^{10}_{00},$$
$$(qra)(qru) b_\alpha p_\beta p_x \gamma_u c_x = f \cdot (FL, f_1)^{10}_{10},$$
$$(qrc)(qru) b_\alpha a_\gamma p_\beta p_x u_x = u_x \cdot (FL, a_\gamma \beta_u c_x b_\alpha)^{10}_{10}.$$

The latter form is reducible in terms of $u_x$, $(FL, f)^{10}_{10}$, $(FL, f_1)^{10}_{10}$, and invariants.

We have now proved that the term $t$ of $\tau$ is reducible.

* Clebsch and Gordan, loc. cit., p. 374.

A different procedure is required in the case of transvectants of the type of

$$\tau = (L, h)^{00}_{02}.$$

This form $\tau$ has the term

$$t = (acu)(\overline{pq}\beta x)(\overline{pq}\gamma x)b_a,$$

and $t$ equals

$$b_a(\overline{pq}\gamma x)\begin{vmatrix}(pqa) & a_\beta & a_x\\ (pqc) & c_\beta & c_x\\ (pqu) & \beta_u & u_x\end{vmatrix}.$$

Expanding the latter determinant we have for one of $t$'s terms

$$t_1 = (pqa)(\overline{pq}\gamma x)\beta_u c_x b_a,$$

and $t_1$ is a term of the transvectant

$$\tau_1 = (L, ff_1)^{00}_{11}.$$

Another term of $\tau_1$ is the reducible form

$$t_2 = (pqa)(\overline{pq}\beta x)b_a\gamma_u c_x = f\cdot(L, f_1)^{00}_{11},$$

and we have, in accordance with the theory of § 1,

$$\begin{aligned} t_1 - t_2 &= (pqa)b_a c_x[(\overline{pq}\gamma x)\beta_u - (\overline{pq}\beta x)\gamma_u]\\ &= (pqa)(pqu)(\beta\gamma x)c_x b_a - (pqa)(pq\overline{\beta\gamma})c_x b_a u_x. \end{aligned}$$

The terms of this last difference are terms of the respective transvectants

$$(L, g)^{00}_{10}, \qquad (L, g)^{00}_{20}\cdot u_x,$$

and these are retained in the above table. Thus $t_1$ is reducible, in the form

$$t_1 = f\cdot(L, f_1)^{00}_{11} + (L, g)^{00}_{10} - u_x\cdot(L, g)^{00}_{20} + \Sigma,$$

where $\Sigma$ is a sum of terms of transvectants $u^\epsilon_x(L, \bar{g})^{ij}_{kl}$, $\bar{g}$ being derived by convolution from $g$ and hence of lower grade than $g$ and thus linear in $f, f_1$, $u_x$ and invariants. The other terms of the expanded $t$, with minus signs changed, are

$$\begin{aligned} (pqa)(\overline{pq}\gamma x)b_a c_\beta u_x &= (L, a_x b_a c_\beta\gamma_u)^{00}_{11}u_x\\ &= (L, a_\gamma\beta_u c_x b_a)^{00}_{11}u_x,\\ (pqc)(\overline{pq}\gamma x)a_x\beta_u b_a &= f_1(L, f)^{00}_{11},\\ (pqu)(\overline{pq}\gamma x)a_\beta b_a c_x &= i_1(L, f)^{00}_{01},\\ (pqu)(\overline{pq}\gamma x)a_x c_\beta b_a &= (L, a_x b_a c_\beta\gamma_u)^{00}_{01},\\ (pqc)(\overline{pq}\gamma x)a_\beta b_a u_x &= i_1\cdot u_x\cdot(L, f)^{00}_{11}. \end{aligned}$$

Hence the term $t$ of $\tau$ is reducible in terms of forms of lower degree, together with forms which are retained in the table and forms which are terms of transvectants for which the grade of $UV$ is less than the grade of $Lh$. Hence $\tau = (L, h)^{00}_{02}$ does not belong to the fundamental system.

The remaining 21 reducible transvectants are as follows:

$(FL, f^2)^{02}_{02}$, $(FL, f_1^2)^{02}_{02}$, $(FL, ff_1)^{02}_{02}$, $(FL, f)^{01}_{01}$, $(FL, f^2)^{02}_{01}$,

$(FL, f^2)^{01}_{02}$, $(FL, f_1)^{01}_{01}$, $(FL, f_1^2)^{02}_{01}$, $(FL, f_1^2)^{01}_{02}$, $(FL, ff_1)^{02}_{01}$,

$(FL, ff_1)^{01}_{02}$, $(F, g)^{01}_{00}$, $(L, h)^{00}_{01}$, $(FL, g)^{01}_{10}$, $(F, g)^{02}_{00}$,

$(FL, g)^{02}_{10}$, $(FL, h)^{10}_{02}$, $(F^2L, fg)^{04}_{01}$, $(FL^2, fh)^{01}_{04}$, $(F^2L, f_1g)^{04}_{01}$,

$(FL^2, f_1h)^{01}_{04}$.

UNIVERSITY OF PENNSYLVANIA

# ALGEBRAIC PROPERTIES OF SELF-ADJOINT SYSTEMS*

BY

DUNHAM JACKSON

The general definition of adjoint systems of boundary conditions associated with ordinary linear differential equations was given by Birkhoff.† In a paper of Bôcher,‡ in which the idea is further developed, there is obtained a condition that a system of the second order be self-adjoint. It is proposed here to extend the discussion of this problem to the case of systems of any order.§ A condition for self-adjointness of the boundary conditions is simply expressed in matrix form, without any requirement that a corresponding property be possessed by the differential equations; the condition gains in symmetry if it is assumed that the differential expression which forms the left-hand member of the given differential equation is itself identical with its adjoint or with the negative of its adjoint.

As a preliminary, it will be well to recall a well-known rule for the combination of matrices, which will be found particularly convenient. Let $\alpha_1$, $\alpha_2$, $\alpha_3$, and $\alpha_4$ be $n$-rowed square matrices. The square matrix of order $2n$ which is made up of the elements of these four, with the elements of $\alpha_1$, arranged in order, in its upper left-hand corner, and the elements of the other matrices correspondingly disposed, may be indicated by the notation

$$\alpha = \begin{pmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{pmatrix}.$$

* Presented to the Society, December 27, 1915.

† G. D. Birkhoff, these Transactions, vol. 9 (1908), pp. 373–395; p. 375.

‡ M. Bôcher, these Transactions, vol. 14 (1913), pp. 403–420; p. 409. Cf. E. Bounitzky, Journal de mathématiques pures et appliquées, 6th series, vol. 5 (1909), pp. 65–125; pp. 112–116; W. Stekloff, *Sur certaines questions d'analyse qui se rattachent à plusieurs problèmes de la physique mathématique*, Mémoires de l'Académie Impériale des Sciences de St.-Pétersbourg, classe physico-mathématique, 8th series, vol. 31 (1913), no. 7; pp. 5–6, 69.

§ Bounitzky, in the paper just cited, pp. 107–116, gives a discussion of the problem, laying stress primarily on the case of a self-adjoint differential equation. His method of treatment is different from that employed here, and in fact bears as little resemblance to it as could well be expected, considering that the author's task in either case, as far as this particular problem is concerned, is merely to give a concise account of algebraic processes which in themselves are perfectly straightforward.

Let $\beta$ be another matrix of the same sort, consisting of the elements of the $n$-rowed matrices $\beta_1, \cdots, \beta_4$. The notation suggests the symbolic product

$$\begin{pmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{pmatrix}\begin{pmatrix} \beta_1 & \beta_2 \\ \beta_3 & \beta_4 \end{pmatrix} = \begin{pmatrix} \alpha_1\beta_1 + \alpha_2\beta_3 & \alpha_1\beta_2 + \alpha_2\beta_4 \\ \alpha_3\beta_1 + \alpha_1\beta_3 & \alpha_3\beta_2 + \alpha_4\beta_4 \end{pmatrix},$$

formed as if the elements indicated were ordinary numbers instead of matrices. By actually writing out the $n$-rowed matrices $\alpha_1\beta_1 + \alpha_2\beta_3$, etc., it is possible to interpret the product as a square matrix of order $2n$; and it is found that this is the same as the product obtained by multiplying together $\alpha$ and $\beta$ as $2n$-rowed matrices in the usual way.

To begin the discussion of the problem in hand, let $L(u)$ stand for the differential expression

$$L(u) = p_0(x)\frac{d^n u}{dx^n} + p_1(x)\frac{d^{n-1}u}{dx^{n-1}} + \cdots + p_n(x)u,$$

where $n$ is any positive integer. The coefficients $p$ are real or complex functions of the real variable $x$, defined in an interval $a \leqq x \leqq b$, and it may be assumed, with more regard for simplicity of statement than for ultimate generality, that they are continuous together with their first $n$ derivatives. With the differential equation

$$L(u) = 0$$

there is supposed given a set of $n$ linearly independent boundary conditions $U_i(u) = 0$, $i = 1, 2, \cdots, n$, in which the general left-hand member is a linear combination, with real or complex constant coefficients, of the $2n$ quantities $u(a), u'(a), \cdots, u^{(n-1)}(a), u(b), u'(b), \cdots, u^{(n-1)}(b)$. Let $n$ of these quantities be chosen in such a way that the determinant of the corresponding coefficients is different from zero. The possibility of such a choice is a consequence of the assumption that the $n$ conditions are linearly independent. The quantities selected will be denoted by $u_1, u_2, \cdots, u_n$, the remaining $n$ quantities by $u_{n+1}, \cdots, u_{2n}$; the order of the subscripts within each group is immaterial. It is most natural to assign the first $n$ subscripts to $u(a), \cdots, u^{(n-1)}(a)$, or to $u(b), \cdots, u^{(n-1)}(b)$, if the corresponding sets of terms are linearly independent, but this will not always be the case.

The differential expression adjoint to $L(u)$ is

$$M(v) = (-1)^n\frac{d^n}{dx^n}(p_0 v) + (-1)^{n-1}\frac{d^{n-1}}{dx^{n-1}}(p_1 v) + \cdots + p_n v.$$

It is connected with $L(u)$ by Lagrange's identity,

$$vL(u) - uM(v) = \frac{d}{dx}P(u, v),$$

where $P(u, v)$ is a bilinear form in $u(x)$, $u'(x)$, $\cdots$, $u^{(n-1)}(x)$; $v(x)$, $v'(x)$, $\cdots$, $v^{(n-1)}(x)$, with coefficients which are functions of $x$. Integrated with regard to $x$ from $a$ to $b$, this identity gives a relation involving the values which the functions just named take on at the ends of the interval. If the values of $v$ and its derivatives at $a$ and $b$ are denoted by $v_1, \cdots, v_{2n}$, in such a way that $u_j$ and $v_j$ refer to derivatives of the same order taken at the same point, the relation becomes

$$\int_a^b [vL(u) - uM(v)]\, dx = \Pi(u, v),$$

in which $\Pi(u, v)$ is a non-singular bilinear form in the $4n$ quantities $u_j$, $v_j$, with constant coefficients.*

The definition of the adjoint boundary conditions proceeds now as follows. In addition to the linear forms $U_1(u), \cdots, U_n(u)$, there are defined $n$ other forms $U_{n+1}(u), \cdots, U_{2n}(u)$, which are arbitrary, except for the restriction that the whole set of $2n$ forms shall be linearly independent. Then $2n$ linear forms $V_1(v), \cdots, V_{2n}(v)$, in the variables $v_j$, are determined so that

$$\Pi(u, v) = \sum_{i=1}^{2n} U_i(u) V_i(v).$$

The last $n$ of these forms, set equal to zero, express the adjoint conditions. It is well known that a different choice of the arbitrary forms $U_{n+1}, \cdots, U_{2n}$, would lead to a set of conditions equivalent collectively to those already obtained, that is, a set of independent linear combinations of them.† It is to be kept in mind throughout that it is the set of equations that is important, and not the individual equation, still less the individual linear form.

The definition just given can be expressed in slightly different words, as follows: The variables $u$ are subjected to a linear transformation $u_i' = U_i(u)$, and then a linear transformation of the other set of variables, $v_i' = V_i(v)$, is determined so that the two transformations together carry over the bilinear form $\Pi(u, v)$ into the normal form $\sum u_i' v_i'$. Let $\Pi$ denote the matrix of the coefficients in the form $\Pi(u, v)$; the double use of the letter $\Pi$, to represent a matrix and to represent an operator, will not cause any confusion. Let $\alpha$ denote the matrix of the coefficients in the *inverse* transformation, by which the $u$'s without accents are expressed in terms of the $u$'s with accents, and let $\beta$ be the corresponding matrix for the $v$'s; the conjugate of $\alpha$ will be de-

* It is to be noticed that the coefficient $p_n$ is without effect on the form of $\Pi$, as the term $p_n uv$ cancels from the difference $vL(u) - uM(v)$. Consequently, if the differential equation with a parameter, $L(u) + \lambda u = 0$, is under discussion, the form $\Pi$ is independent of the parameter, and the same is true of the adjoint boundary conditions, which depend on $\Pi$ for their definition.

† Birkhoff, loc. cit., p. 375; Bôcher, loc. cit., p. 405.

noted by $\alpha'$, and, in general, conjugate matrices will be indicated by accents. Finally, let $I$ denote the unit matrix of order $2n$, having unity for each element in the principal diagonal, and zero everywhere else. Then the relation connecting these matrices is simply the equation*

$$\alpha' \Pi \beta = I. \tag{1}$$

It is to be noticed that not $\alpha$, but $\alpha^{-1}$, is to be regarded as given directly, and $\beta^{-1}$ is the matrix that is ultimately desired. An expression for the latter is found by multiplying both members of (1) on the right by $\beta^{-1}$:

$$\beta^{-1} = \alpha' \Pi.$$

Hence the problem to be solved may be formulated thus:

*Under what circumstances will the last n rows of* $\alpha' \Pi$ *be independent linear combinations of the first n rows of* $\alpha^{-1}$*?*

Let $\alpha_1$ stand for the square matrix of the coefficients of the first $n$ variables $u_j$ in the $n$ given forms $U_i$, and $\alpha_2$ for the matrix of the remaining coefficients in these forms. Since the notation has been supposed chosen so that $\alpha_1$ is non-singular, the $2n$ forms $U_1, \cdots, U_{2n}$ will be linearly independent if the last $n$ of them are set equal to $u_{n+1}, \cdots, u_{2n}$ respectively. Then the matrix $\alpha^{-1}$ has the form

$$\alpha^{-1} = \begin{pmatrix} \alpha_1 & \alpha_2 \\ 0 & I \end{pmatrix};$$

here $I$ denotes the unit matrix of order $n$ (the use of the same letter to represent unit matrices of different orders will not cause any confusion) and 0 stands for a matrix composed entirely of zeros. It may be verified at once that

$$\begin{pmatrix} \alpha_1 & \alpha_2 \\ 0 & I \end{pmatrix} \begin{pmatrix} \alpha_1^{-1} & -\alpha_1^{-1}\alpha_2 \\ 0 & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix},$$

from which it follows that the second factor on the left is the matrix $\alpha$. The product $\alpha_1^{-1}\alpha_2$ will be denoted for brevity by the single letter $\delta$. It will be convenient also to indicate the conjugate of the reciprocal of a non-singular matrix by a double accent, so that $(\alpha_1^{-1})' = (\alpha_1')^{-1} = \alpha_1''$, and $\delta' = \alpha_2' \alpha_1''$. Then we have for the conjugate of $\alpha$ the representation

$$\alpha' = \begin{pmatrix} \alpha_1'' & 0 \\ -\delta' & I \end{pmatrix}.$$

The matrix of the bilinear form will be represented for the present simply by the notation

$$\Pi = \begin{pmatrix} \pi_1 & \pi_3 \\ \pi_4 & \pi_2 \end{pmatrix},$$

* Cf., e. g., Bôcher, *Introduction to Higher Algebra*, p. 115.

which is applicable to any square matrix of order $2n$; with regard to the special form of $\Pi$, more will be said later. The expression for $\beta^{-1}$ is

$$\alpha' \Pi = \begin{pmatrix} \alpha_1'' \pi_1 & \alpha_1'' \pi_3 \\ -\delta' \pi_1 + \pi_4 & -\delta' \pi_3 + \pi_2 \end{pmatrix}.$$

Now, since $\alpha_1$ is non-singular, it is possible to express any sets of $n$ quantities each, in particular, the rows of the matrix standing in the lower left-hand corner of $\alpha' \Pi$, as linear combinations of the rows of $\alpha_1$; and the coefficients by which this is accomplished are uniquely determined. The question is, then, whether the corresponding linear combinations of the rows of $\alpha_2$ are identical respectively with the rows of the matrix in the lower right-hand corner of $\alpha' \Pi$. It will be seen presently that if this is the case, the combinations in question will necessarily be independent.

In general terms, if $\eta$ is a matrix of order $n$, and there are formed $n$ linear combinations of its rows, the coefficients in these combinations forming the successive rows of a matrix $\gamma$, the resulting sets of $n$ quantities each are respectively the rows of the product matrix $\gamma\eta$. In the present instance, $\gamma$ is to be determined so that

$$\gamma\alpha_1 = -\delta' \pi_1 + \pi_4, \tag{2}$$

and the condition for a self-adjoint system is that the matrix $\gamma$ so defined satisfy the further relation

$$\gamma\alpha_2 = -\delta' \pi_3 + \pi_2. \tag{3}$$

If it does, then

$$\alpha' \Pi = \begin{pmatrix} \alpha_1'' \pi_1 & \alpha_1'' \pi_3 \\ \gamma\alpha_1 & \gamma\alpha_2 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & \gamma \end{pmatrix} \begin{pmatrix} \alpha_1'' \pi_1 & \alpha_1'' \pi_2 \\ \alpha_1 & \alpha_2 \end{pmatrix}.$$

If $\gamma$ were singular, the matrix $\alpha' \Pi$ would be singular, which is certainly not the case; it follows that the determinant of $\gamma$ is necessarily different from zero. Let (2) be solved for $\gamma$, and the result substituted in (3). It is found that

$$\gamma = -\delta' \pi_1 \alpha_1^{-1} + \pi_4 \alpha_1^{-1}, \qquad -\delta' \pi_1 \delta + \pi_4 \delta = -\delta' \pi_3 + \pi_2. \tag{4}$$

In the last equation, $\delta$ has been substituted in the left-hand member for its equal, $\alpha_1^{-1} \alpha_2$. The relation (2), defining $\gamma$, may be regarded as valid in any case. Then either of the relations (3), (4), is a consequence of the other. The conclusion may be stated as follows:

THEOREM. *A necessary and sufficient condition that the given set of boundary conditions be self-adjoint, is that**

$$\delta' \pi_1 \delta - \pi_4 \delta = \delta' \pi_3 - \pi_2. \tag{5}$$

* Professor W. A. Hurwitz points out that this condition can be written concisely in the form $\Delta' \Pi \Delta = 0$, where

$$\Delta = \begin{pmatrix} 0 & -\delta \\ 0 & I \end{pmatrix}.$$

In the cases that are of most interest, the form of the condition can be simplified. Suppose first that in the $n$ given boundary conditions, the sets of terms corresponding to one end of the interval are linearly independent, so that the values of $u$ and its derivatives at this point may be represented by $u_1, \cdots, u_n$. Then, as no single term of $\Pi(u, v)$ is concerned with both ends of the interval, the subscripts in each term that actually appears will be both less than $n+1$ or both greater than $n$, and the matrices previously denoted by $\pi_3$ and $\pi_4$ will be composed entirely of zeros.* If $\delta$ is replaced by the alternative expression $\alpha_1^{-1}\alpha_2$, the condition for self-adjointness reads as follows:

$$\alpha_2'\alpha_1''\pi_1\alpha_1^{-1}\alpha_2 = -\pi_2.$$

It appears that in this case $\alpha_2$ as well as $\alpha_1$ must be non-singular, since $\pi_2$ is non-singular. On multiplication before and after by $\alpha_2''$ and $\alpha_2^{-1}$ respectively, the condition takes the form

$$\alpha_1''\pi_1\alpha_1^{-1} = -\alpha_2''\pi_2\alpha_2^{-1},$$

the meaning of which can be expressed in still different language. There is no loss of generality in assuming, for definiteness, that $\alpha_1$ is the matrix of the coefficients belonging to the point $b$. More precisely, it may be assumed that

$$u_1 = u(b), \quad u_2 = u'(b), \quad \cdots, \quad u_n = u^{(n-1)}(b),$$

$$u_{n+1} = u(a), \quad u_{n+2} = u'(a), \quad \cdots, \quad u_{2n} = u^{(n-1)}(a),$$

while $v_1, \cdots, v_{2n}$ have a corresponding interpretation. Let the value of $P(u, v)$ for $x = b$ be denoted by $\Pi_1(u, v)$, and its value for $x = a$, by $\Pi_2(u, v)$, so that

$$\Pi(u, v) = \Pi_1(u, v) - \Pi_2(u, v).$$

The expressions $\Pi_1$ and $\Pi_2$ are bilinear forms in $u_1, \cdots, u_n; v_1, \cdots, v_n$, and $u_{n+1}, \cdots, u_{2n}; v_{n+1}, \cdots, v_{2n}$, respectively, and their matrices are $\pi_1$ and $-\pi_2$. Let each of the $n$ given linear forms $U_i(u)$ be regarded as the sum of a form $U_{i1}(u)$, involving only the first $n$ of the variables $u_j$, and a form $U_{i2}(u)$, involving only the last $n$ variables. If $\Pi_1(u, v)$ is written as a bilinear form in $U_{i1}(u)$ and $U_{i1}(v)$, $i = 1, 2, \cdots, n$, the coefficients in the representation have the matrix $\alpha_1''\pi_1\alpha_1^{-1}$. A corresponding remark applies to $\Pi_2(u, v)$. It follows that

*If the forms $U_{i1}(u)$ are linearly independent, the condition for a self-adjoint system is that the coefficients by means of which $\Pi_1(u, v)$ is expressed in terms of the linear forms $U_{i1}(u)$, $U_{i1}(v)$, be the same as those by which $\Pi_2(u, v)$ is expressed in terms of the forms $U_{i2}(u)$, $U_{i2}(v)$.*

* The fact that $\pi_1$ and $\pi_2$ also contain a considerable number of zeros need not be taken into account.

Now let the restriction on the choice of the first $n$ variables be removed, but let it be supposed that when the given differential expression $L(u)$ and its adjoint are written with the same function $u$ as argument, the latter expression is identical with $L(u)$ or with the negative of $L(u)$, in other words, that $L(u)$ is self-adjoint or anti-self-adjoint. The statement of the theorem is again capable of some simplification, for the reason now that the bilinear form $P(u, v)$ is skew-symmetric in the one case and symmetric in the other,* and the same thing consequently is true of the matrix $\Pi$. More generally, it suffices for the symmetry of $P$ and $\Pi$ that $L(u)$ differ only in the term $p_n u$ from an expression which is anti-self-adjoint,† since, as has already been remarked, this term has no effect on the form of $P$.

Let the condition (5) be written in the equivalent form

$$\delta' \pi_1 \delta - 2\pi_4 \delta + \pi_2 = -\delta' \pi_1 \delta + 2\delta' \pi_3 - \pi_2. \tag{6}$$

If $\Pi$ is skew-symmetric, then $\pi_1' = -\pi_1$, $\pi_2' = -\pi_2$, and $\pi_4' = -\pi_3$, and it appears that the matrix which forms the right-hand member of (6) is the conjugate of the matrix on the left. If $\Pi$ is symmetric, each member of (6) is the negative of the conjugate of the other. That is:

*If the differential expression $L(u)$ is self-adjoint, the condition that the boundary conditions be self-adjoint is that the matrix*

$$\delta' \pi_1 \delta - 2\pi_4 \delta + \pi_2$$

*be symmetric; if $L(u)$ is an anti-self-adjoint expression, or differs from such an expression only in the term of order zero, the condition is that the matrix just written down be skew-symmetric.*

This form of statement remains valid, of course, under the earlier hypothesis that $\pi_4 = 0$; but in that case the condition simply resumes the form

$$\delta' \pi_1 \delta + \pi_2 = 0,$$

since this matrix is required to be symmetric and skew-symmetric at the same time.‡

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

* Cf. L. Schlesinger, *Handbuch der Theorie der linearen Differentialgleichungen*, vol. 1, pp. 70–71, 73–74.

† This observation is not trivial, as the corresponding statement with regard to self-adjoint expressions would be; for example, the expression $xu' + \frac{1}{2}u$ is anti-self-adjoint, while the expression $xu'$ is not.

‡ It is necessarily skew-symmetric if $L(u)$ is self-adjoint, without regard to the boundary conditions, and becomes symmetric if the latter are self-adjoint, and similarly in the anti-self-adjoint case.

# ON THE SECOND DERIVATIVES OF AN EXTREMAL-INTEGRAL WITH AN APPLICATION TO A PROBLEM WITH VARIABLE END POINTS*

BY

ARNOLD DRESDEN

## INTRODUCTION

In a paper published in volume 9 of these Transactions, formulæ were obtained for the second derivatives of the extremal-integral arising in the problem of minimizing or maximizing the integral $\int F(x, y, x', y')\,dt$. In these formulæ the derivatives were expressed in terms of particular solutions of the Weierstrass form of Jacobi's differential equation.† They were used in the same paper to obtain further necessary conditions for a minimum of the above mentioned integral, when one or both endpoints were variable along a curve,‡ or when curves with discontinuous slopes were admitted as solutions of the problem. It is the author's aim to use the same general method in analogous problems involving integrals of a more general type, and also in the cases in which the unknown functions are further conditioned by differential and algebraic equations. As a first step in this direction, we obtain in the present paper the second derivatives of the extremal-integral arising in the theory of the integral

$$\int f(x, y_1, \cdots, y_n; y_1', \cdots, y_n')\,dx \qquad (y_i' = dy_i/dx), \tag{1}$$

and express them in terms of particular solutions of the differential equations which correspond to Jacobi's equations for the simpler problem. The formulæ for these derivatives are then used for finding further necessary conditions to be satisfied by a curve $C$ in the space of $n+1$ dimensions which is to furnish a

* The contents of sections 1 and 2 of this paper are contained in the paper presented to the Society at the Chicago meeting of December, 1914, an abstract of which appeared in the Proceedings of the National Academy of Sciences for March, 1914.

† Compare Bolza, *Vorlesungen über Variationsrechnung*, p. 306–313 (this book will be referred to in the sequel by the author's name only), and Dresden, these Transactions, vol. 9 (1908), pp. 467 and 472.

‡ The conditions here referred to had been found before by Bliss, these Transactions, vol. 3 (1902), p. 136, and Mathematische Annalen, vol. 58 (1903), p. 70.

minimum for the integral (1), when the initial point of $C$ is allowed to move along a fixed curve $L$.*

A number of familiar facts concerning the integral (1) which will be needed in the development of our problem have been collected in section 1. In section 2 the formulæ for the second derivatives of the extremal integral are obtained, and in section 3 these formulæ are applied to the problem with variable initial point.

Concerning the function $f$, which appears in the integral (1), we assume that it is of class $C'''$ with respect to all of its arguments; the functions $y(x)$ which are admitted as solutions of our problem must be of class $C''$. We shall moreover use the following abbreviated notations:

$$f_0 = \partial f/\partial x, \qquad f_i = \partial f/\partial y_i, \qquad f_{n+i} = \partial f/\partial y_i',$$

and similar notations will be used for the second derivatives. A point in $n+1$ dimensional space of coördinates $x, y_1, \cdots, y_n$ will be designated by $(x; y)$, and similar notations will be used for sets of numbers which can be interpreted as coördinates of points in spaces of higher dimensionality. Finally the notation $f^{(i)}$ will be used to indicate that the function $f(x; y; y')$ is to be taken with the arguments $(x_i; y_i; y_i')$.

## 1. Preliminary developments

Any set of functions $y_1(x), \cdots, y_n(x)$ which is to furnish a minimum or a maximum for the integral (1) must satisfy the system of differential equations

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx}\frac{\partial f}{\partial y_i'} = 0 \qquad (i = 1, \cdots, n), \tag{2}$$

known as Euler's differential equations.† Such a set of functions defines a curve in the space of $n+1$ dimensions, called an extremal for the problem; we shall use repeatedly geometric terminology consistent with this interpretation of the set of functions $y_i(x)$. From the general existence theorems for differential equations,‡ we know that the general solution of the system of differential equations (2) is of the form

$$y_i = g_i(x; \alpha_1, \cdots, \alpha_n; \beta_1, \cdots, \beta_n) \qquad (i = 1, \cdots, n), \tag{3}$$

the functions $g_i$ being of class $C'''$ with respect to all of their arguments and $\alpha_1, \cdots, \beta_n$ being $2n$ arbitrary constants.

* For the case $n = 2$, this problem has been treated by Miss M. B. White in her paper, *The dependence of focal points upon curvature for problems of the calculus of variations in space*, these Transactions, vol. 13 (1912), p. 175.

† See Bolza, p. 542.

‡ Ibid., pp. 175–179.

We suppose now that we have found an extremal passing through the two points $A(a_0; a)$ and $B(b_0; b)$ and satisfying the conditions of Clebsch and Mayer.* The second variation of the integral (1) can then be put into the form

$$\delta^2 I = \frac{\epsilon^2}{2} \int_{a_0}^{b_0} \sum_{ik} (P_{ik}\, \eta_i\, \eta_k + 2Q_{ik}\, \eta_i\, \eta_k' + R_{ik}\, \eta_i'\, \eta_k')\, dx,$$

where $\eta_1, \cdots, \eta_n$ are arbitrary functions of $x$ of class $C'$ vanishing for $x = a_0$ and for $x = b_0$, and where $P_{ik} = f_{ik}$, $Q_{ik} = f_{i,\, n+k}$ and $R_{ik} = f_{n+i,\, n+k}$, these functions being taken along the given extremal. If the integrand which appears in this formula for the second variation be denoted by $2\Omega$, and if we write

$$\psi_i(\eta) = \frac{\partial \Omega}{\partial \eta_i} - \frac{d}{dx} \frac{\partial \Omega}{\partial \eta_i'},$$

we obtain, by means of Euler's theorem on homogeneous functions and by the familiar integration by parts, the following result:

$$\delta^2 I = \frac{\epsilon^2}{2} \int_{a_0}^{b_0} \sum_i \eta_i\, \psi_i\, dx.$$

The system of equations

$$\psi_i(\eta) = 0, \qquad (i = 1, \cdots, n), \tag{4}$$

is a system of linear differential equations of the second order in the unknown functions $\eta_i$; it is a self-adjoint system† and is called, after Von Escherich, the accessory system of differential equations;‡ we observe moreover that it is the direct analogue of Jacobi's equation for the simplest problem of the calculus of variations. It follows further§ that a fundamental system of $2n$ solutions of the equations (4) may be obtained by differentiating the functions $g_i$, which appear in the equations (3), with respect to each of the parameters $\alpha_1, \cdots, \beta_n$ and then replacing these by $\alpha_{10}, \cdots, \beta_{n0}$, the latter being the values which furnish the given extremal if put in place of $\alpha_1, \cdots, \beta_n$ in equations (3). We obtain in this way the following $2n$ systems of solutions of equations (4):

$$\partial g_1/\partial \alpha_i, \cdots, \partial g_n/\partial \alpha_i; \qquad \partial g_1/\partial \beta_i, \cdots, \partial g_n/\partial \beta_i, \quad (i = 1, \cdots, n), \tag{5}$$

the arguments of these functions being $x$; $\alpha_{10}, \cdots, \beta_{n0}$.

We consider now two points $P_1(x_1; y_1)$ and $P_2(x_2; y_2)$ in the neighborhood of $A(a_0; a)$ and $B(b_0; b)$, respectively. Since the extremal $AB$ satisfies the Mayer condition, we can conclude that it is possible to construct an

* Bolza, p. 608 and p. 619.

† Hadamard, *Leçons sur le Calcul des Variations*, p. 319.

‡ Bolza, p. 622; further references are given there.

§ Bolza, p. 623 and Hadamard, loc. cit., pp. 336–339.

unique extremal through the points $P_1$ and $P_2$, i. e., that it is possible to solve the equations

$$(6) \qquad y_{1i} = g_i(x_1; \alpha_1, \cdots, \beta_n), \qquad y_{2i} = g_i(x_2; \alpha_1, \cdots, \beta_n)$$

uniquely for $\alpha_i$ and $\beta_i$ in terms of $x_1$, $x_2$, $y_{1j}$ and $y_{2j}$.* We denote the solutions of these equations by:

$$(7) \quad \alpha_i = A_i(x_1, x_2; y_{11}, \cdots, y_{2n}), \qquad \beta_i = B_i(x_1, x_2; y_{11}, \cdots, y_{2n}).$$

Substitution of equations (7) in equations (3) furnishes the following analytical representation for the extremal through the points $P_1$ and $P_2$:

$$(8) \qquad y_i = g_i\big(x; A_1(x_1, \cdots, y_{2n}), \cdots, B_n(x_1, \cdots, y_{2n})\big) \equiv Y_i(x; x_1, x_2; y_{11}, \cdots, y_{2n}).$$

Upon substituting from equations (7) in equations (6), we see that the functions $Y_i$, which are defined by equations (8), satisfy the following initial conditions:

$$(9) \quad Y_i(x_1; x_1, x_2; y_{11}, \cdots, y_{2n}) = y_{1i}, \qquad Y_i(x_2; x_1, x_2, y_{11}, \cdots, y_{2n}) = y_{2i}.$$

Differentiation of these equations with respect to $x_1, x_2, y_{11}, \cdots, y_{2n}$ furnishes furthermore the following relations, which will be of use in the sequel:

$$(10) \quad \begin{aligned} &\partial Y_i/\partial x_1|_1 + Y_i'(x_1) = 0, \qquad \partial Y_i/\partial x_2|_2 + Y_i'(x_2) = 0, \\ &\partial Y_i/\partial y_{2j}|_1 = 0, \qquad \partial Y_i/\partial x_2|_1 = 0, \qquad \partial Y_i/\partial y_{1j}|_1 = \delta_{ij}, \\ &\partial Y_i/\partial x_1|_2 = 0, \qquad \partial Y_i/\partial y_{1j}|_2 = 0, \qquad \partial Y_i/\partial y_{2j}|_2 = \delta_{ij}, \end{aligned}$$

or, expressed in different form,

$$g_i'(x_1) + \sum_j \left\{ \frac{\partial g_i}{\partial \alpha_j}\frac{\partial A_j}{\partial x_1} + \frac{\partial g_i}{\partial \beta_j}\frac{\partial B_j}{\partial x_1} \right\}\Big|_1 = 0,$$

$$g_i'(x_2) + \sum_j \left\{ \frac{\partial g_i}{\partial \alpha_j}\frac{\partial A_j}{\partial x_2} + \frac{\partial g_i}{\partial \beta_j}\frac{\partial B_j}{\partial x_2} \right\}\Big|_2 = 0,$$

$$(11) \quad \begin{aligned} &\sum_j \left\{ \frac{\partial g_i}{\partial \alpha_j}\frac{\partial A_j}{\partial x_1} + \frac{\partial g_i}{\partial \beta_j}\frac{\partial B_j}{\partial x_1} \right\}\Big|_2 = 0, \qquad \sum_j \left\{ \frac{\partial g_i}{\partial \alpha_j}\frac{\partial A_j}{\partial x_2} + \frac{\partial g_i}{\partial \beta_j}\frac{\partial B_j}{\partial x_2} \right\}\Big|_1 = 0, \\ &\sum_j \left\{ \frac{\partial g_i}{\partial \alpha_j}\frac{\partial A_j}{\partial y_{1h}} + \frac{\partial g_i}{\partial \beta_j}\frac{\partial B_j}{\partial y_{1h}} \right\}\Big|_1 = \delta_{ih}, \qquad \sum_j \left\{ \frac{\partial g_i}{\partial \alpha_j}\frac{\partial A_j}{\partial y_{2h}} + \frac{\partial g_i}{\partial \beta_j}\frac{\partial B_j}{\partial y_{2h}} \right\}\Big|_1 = 0, \\ &\sum_j \left\{ \frac{\partial g_i}{\partial \alpha_j}\frac{\partial A_j}{\partial y_{1h}} + \frac{\partial g_i}{\partial \beta_j}\frac{\partial B_j}{\partial y_{1h}} \right\}\Big|_2 = 0, \qquad \sum_j \left\{ \frac{\partial g_i}{\partial \alpha_j}\frac{\partial A_j}{\partial y_{2h}} + \frac{\partial g_i}{\partial \beta_j}\frac{\partial B_j}{\partial y_{2h}} \right\}\Big|_2 = \delta_{ih}, \end{aligned}$$

$$(i = 1, \cdots, n;\ h = 1, \cdots, n).$$

* Bolza, p. 597.

We compute now the integral (1) along the extremal $P_1P_2$ by substituting the functions $Y_1, \cdots, Y_n$ and their derivatives with respect to $x$, for $y_1, \cdots, y_n$ and $y_1', \cdots, y_n'$ respectively in the integral (1). This integral becomes then a function of the $2n + 2$ coördinates of the points $P_1$ and $P_2$;* it is known as the "extremal-integral":

$$I(x_1, x_2, y_{11}, \cdots, y_{2n}) = \int_{x_1}^{x_2} f(x; Y_1, \cdots, Y_n; Y_1', \cdots, Y_n')\, dx.$$

The first derivatives of this function I with respect to its different arguments were first given by Hamilton,* and we shall use them as a basis for the computation of the second derivatives; we have the following formulæ:

$$\begin{aligned} \partial I/\partial x_1 &= -f(x_1; y_1; y_1') + \sum_i f_{n+i}(x_1; y_1; y_1')\, y_{1i}', \\ \partial I/\partial x_2 &= f(x_2; y_2; y_2') - \sum_i f_{n+i}(x_2; y_2; y_2')\, y_{2i}', \\ \partial I/\partial y_{1j} &= -f_{n+j}(x_1; y_1; y_1'), \qquad \partial I/\partial y_{2j} = f_{n+j}(x_2; y_2; y_2'), \end{aligned} \tag{12}$$

where $y_1, y_1'$ are abbreviations for the sets $(y_{11}, \cdots, y_{1n})$ and $(y_{11}', \cdots, y_{1n}')$ respectively.

## 2. The second derivatives of the extremal-integral

For the determination of the second derivatives of the function $I$ it is necessary to know the derivatives with respect to $x_1, \cdots, y_{2n}$ of the quantities $y_{11}', \cdots, y_{1n}'$ and $y_{21}', \cdots, y_{2n}'$ which determine the slope of the extremal $P_1P_2$, at its end-points. If we denote by $z$ any one of the coördinates of the points $P_1$ and $P_2$, we obtain immediately, upon differentiation of equations (8), the result,

$$\frac{\partial Y_i}{\partial z} = \sum_j \left( \frac{\partial g_i}{\partial \alpha_j} \frac{\partial A_j}{\partial z} + \frac{\partial g_i}{\partial \beta_j} \frac{\partial B_j}{\partial z} \right), \tag{13}$$

from which the desired derivatives may be obtained by differentiation with respect to $x$. Furthermore equations (11) enable us to determine expressions for $\partial A_j/\partial z$ and $\partial B_j/\partial z$; for we have there a set of $2n$ linear non-homogeneous equations for each of the $2n + 2$ sets of $2n$ unknowns which we get if we let $z$ assume in succession each of its values. These sets of equations all possess finite unique solutions provided Mayer's condition is satisfied along the extremal $P_1P_2$. That this will indeed be the case, if the points $P_1$ and $P_2$ are taken sufficiently near to the points $A$ and $B$, respectively, follows from our conditions on the extremal $AB$ and from the continuity of Mayer's determinant:†

* Bolza, p. 599.

† Bolza, p. 611.

$$(14)\qquad \Delta(x, x_1) = \begin{vmatrix} \frac{\partial g_1}{\partial \alpha_1} & \cdots & \frac{\partial g_1}{\partial \alpha_n} & \frac{\partial g_1}{\partial \beta_1} & \cdots & \frac{\partial g_1}{\partial \beta_n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial g_n}{\partial \alpha_1} & \cdots & \frac{\partial g_n}{\partial \alpha_n} & \frac{\partial g_n}{\partial \beta_1} & \cdots & \frac{\partial g_n}{\partial \beta_n} \\ \frac{\partial g_1}{\partial \alpha_1}\Big|_1 & \cdots & \frac{\partial g_1}{\partial \alpha_n}\Big|_1 & \frac{\partial g_1}{\partial \beta_1}\Big|_1 & \cdots & \frac{\partial g_1}{\partial \beta_n}\Big|_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial g_n}{\partial \alpha_1}\Big|_1 & \cdots & \frac{\partial g_n}{\partial \alpha_n}\Big|_1 & \frac{\partial g_n}{\partial \beta_1}\Big|_1 & \cdots & \frac{\partial g_n}{\partial \beta_n}\Big|_1 \end{vmatrix}.$$

We denote now by $\Delta(x, x_2)$ the determinant obtained from $\Delta(x, x_1)$ by replacing the substitution symbol $|_1$ in the last $n$ rows by $|_2$; by $\Delta(x_1, x_2)$ the result of replacing the first $n$ rows of $\Delta(x, x_2)$ by the last $n$ rows of $\Delta(x, x_1)$; by $\Delta_{jk}(x; x_1, x_2)$ the result of replacing the $j$th row of $\Delta(x_1, x_2)$ by the $k$th row of $\Delta(x, x_1)$; and by $\nabla_{jk}(x; x_1, x_2)$ the result of replacing the $(n+j)$th row of $\Delta(x_1, x_2)$ by the $k$th row of $\Delta(x, x_2)$.

With this notation we can write down the results of solving equations (11) for $\partial A_j/\partial z$ and $\partial B_j/\partial z$; substitution of them in equations (13) leads to the following formulæ:

$$\frac{\partial Y_k}{\partial y_{1j}} = \frac{\Delta_{jk}(x; x_1, x_2)}{\Delta(x_1, x_2)}, \qquad \frac{\partial Y_k}{\partial x_1} = -\sum_i g_i'(x_1)\frac{\Delta_{ik}(x; x_1, x_2)}{\Delta(x_1, x_2)},$$

$$\frac{\partial Y_k}{\partial y_{2j}} = \frac{\nabla_{jk}(x; x_1, x_2)}{\Delta(x_1, x_2)}, \qquad \frac{\partial Y_k}{\partial x_2} = -\sum_i g_i'(x_2)\frac{\nabla_{ik}(x; x_1, x_2)}{\Delta(x_1, x_2)}.$$

From these we obtain by direct differentiation the derivatives with respect to $x_1, \cdots, y_{2n}$ of the quantities $y_{11}', \cdots, y_{2n}'$, namely:

$$(15)\qquad \frac{\partial^2 Y_k}{\partial x \partial y_{1j}} = \frac{\Delta_{jk}'(x; x_1, x_2)}{\Delta(x_1; x_2)}, \qquad \frac{\partial^2 Y_k}{\partial x \partial x_1} = -\sum_i g_i'(x_1)\frac{\Delta_{ik}'(x; x_1, x_2)}{\Delta(x_1, x_2)},$$

$$\frac{\partial^2 Y_k}{\partial x \partial y_{2j}} = \frac{\nabla_{jk}'(x; x_1, x_2)}{\Delta(x_1, x_2)}, \qquad \frac{\partial^2 Y_k}{\partial x \partial x_2} = -\sum_i g_i'(x_2)\frac{\nabla_{ik}'(x; x_1, x_2)}{\Delta(x_1, x_2)}.$$

From the definition of the functions $\Delta_{jk}$ and $\nabla_{jk}$ combined with the properties of the functions (5), it follows that the functions

$$u_{jk} = \Delta_{jk}(x; x_1, x_2)/\Delta(x_1, x_2), \qquad v_{jk} = \nabla_{jk}(x; x_1, x_2)/\Delta(x_1, x_2)$$

furnish for every fixed $j$ a solution of the accessory system of differential equations, and that they satisfy the following initial conditions:

$$(16)\qquad u_{jk}(x_1) = \delta_{jk}, \qquad u_{jk}(x_2) = 0; \qquad v_{jk}(x_1) = 0, \qquad v_{jk}(x_2) = \delta_{jk}.$$

These initial conditions show moreover that each of the systems of functions

$u_{jk}$ and $v_{jk}$ constitute, in the terminology of Von Escherich, a "conjugate system" of solutions of the accessory system of differential equations.*

We can now express the second derivatives of the extremal-integral in terms of these solutions of the accessory system of differential equations. Differentiation of equations (12) yields, by the use of equations (10) and (15), the following theorem:

THEOREM I. *The second derivatives of the extremal integral for the integral*

$$\int f(x; y_1, \cdots, y_n; y'_1, \cdots, y'_n)\, dx$$

*are given by the formulæ*

$$I_{1,1} = -f^{(1)}_0 + \sum_i f^{(1)}_{0,n+i}\, y'_{1i} + \sum_{ij} f^{(1)}_{n+i,n+j}\, y'_{1i}\, y'_{1j} - \sum_{ijk} f^{(1)}_{n+i,n+j}\, y'_{1i}\, y'_{1k}\, u'_{kj}(x_1),$$

$$I_{2,2} = f^{(2)}_0 - \sum_i f^{(2)}_{0,n+i}\, y'_{2i} - \sum_{ij} f^{(2)}_{n+i,n+j}\, y'_{2i}\, y'_{2j} + \sum_{ijk} f^{(2)}_{n+i,n+j}\, y'_{2i}\, y'_{2k}\, v'_{kj}(x_2),$$

$$I_{1,2} = -\sum_{ijk} f^{(1)}_{n+i,n+j}\, y'_{1i}\, y'_{2k}\, v'_{kj}(x_1) = \sum_{ijk} f^{(2)}_{n+i,n+j}\, y'_{2i}\, y'_{1k}\, u'_{kj}(x_2),$$

$$I_{1,1i} = -f^{(1)}_i + \sum_j f^{(1)}_{n+j,i}\, y'_{1j} + \sum_{jk} f^{(1)}_{n+j,n+k}\, y'_{1j}\, u'_{ik}(x_1),$$

$$= -f^{(1)}_{0,n+i} - \sum_j f^{(1)}_{n+i,n+j}\, y'_{1j} + \sum_{jk} f^{(1)}_{n+i,n+j}\, y'_{1k}\, u'_{kj}(x_1),$$

$$I_{1,2i} = \sum_{jk} f^{(1)}_{n+j,n+k}\, y'_{1j}\, v'_{ik}(x_1) = -\sum_{jk} f^{(2)}_{n+i,n+j}\, y'_{1k}\, u'_{kj}(x_2), \tag{17}$$

$$I_{2,1i} = -\sum_{jk} f^{(2)}_{n+j,n+k}\, y'_{2j}\, u'_{ik}(x_2) = \sum_{jk} f^{(1)}_{n+i,n+j}\, y'_{2k}\, v'_{kj}(x_1),$$

$$I_{2,2i} = f^{(2)}_i - \sum_j f^{(2)}_{n+j,i}\, y'_{2j} - \sum_{jk} f^{(2)}_{n+j,n+k}\, y'_{2j}\, v'_{ik}(x_2),$$

$$= f^{(2)}_{0,n+i} + \sum_j f^{(2)}_{n+i,n+j}\, y''_{2j} - \sum_{jk} f^{(2)}_{n+i,n+j}\, y'_{2k}\, v'_{kj}(x_2),$$

$$I_{1i,1j} = -\sum_k f^{(1)}_{n+i,n+k}\, u'_{jk}(x_1) - f^{(1)}_{n+i,j}$$

$$= -\sum_k f^{(1)}_{n+j,n+k}\, u'_{ik}(x_1) - f^{(1)}_{n+j,i},$$

$$I_{1i,2j} = -\sum_k f^{(1)}_{n+i,n+k}\, v'_{jk}(x_1) = \sum_k f^{(2)}_{n+j,n+k}\, u'_{ik}(x_2),$$

$$I_{2i,2j} = f^{(2)}_{n+i,j} + \sum_k f^{(2)}_{n+i,n+k}\, v'_{jk}(x_2) = f^{(2)}_{n+j,i} + \sum_k f^{(2)}_{n+j,n+k}\, v'_{ik}(x_2),$$

* See Bolza, p. 627; Hadamard, loc. cit., p. 339; also von Escherich, Wiener Berichte, vol. 107 (1898), p. 1222.

*where the functions $u_{ik}$ and $v_{ik}$ are the solutions of the accessory system of differential equations* (4) *which are characterized by the initial conditions* (16). The subscripts of I specify the variables with respect to which the differentiations are made, single subscripts designating the variable $x$, while pairs of subscripts designate the $y$'s. Thus $I_{2,1i} = \partial^2 I/\partial x_2 \partial y_{1i}$.

From the continuity restrictions imposed upon the function $f$, it follows that the results of differentiating the extremal integral should be independent of the order of differentiation and hence that the two different expressions given for certain of the derivatives in formulæ (17) should be equal to one another. This fact leads to eight relations between the expressions given in theorem I. It is important for certain applications and as a check upon the accuracy of the formulæ themselves, that these relations should be derived directly from the properties of the functions $u_{ik}$ and $v_{ik}$. We shall denote by $(A)$, $(B)$, $(C)$, $(D)$, $(E)$, $(F)$, $(G)$, and $(H)$ the relations resulting from setting equal to one another the two expressions occurring in each of formulæ $(17_3)$ to $(17_{10})$ respectively.

In the first place we observe that certain of these relations follow from others; the remaining ones will then be proved to hold as a consequence of the properties of the functions $u_{ik}$ and $v_{ik}$. If in $(G)$ we interchange the indices $i$ and $j$, then multiply through by $y'_{1j}$ and sum with respect to $j$, and finally interchange the indices $j$ and $k$ in the second sum, we obtain $(C)$. If $(G)$ is multiplied through by $y'_{2j}$ and then summed with respect to $j$, and if finally the summation indices in the first sum are interchanged, we obtain $(D)$. Next, we multiply $(D)$ through by $y'_{1i}$, sum with respect to $i$ and replace the indices $i, j$, and $k$ in the second sum by $k, j$, and $i$, respectively; in this way we obtain $(A)$.*

The functions $y_i(x)$ which occur in our formulæ are solutions of Euler's differential equations (2). Substituting these functions in the differential equations and carrying out the differentiation in the second member, we obtain the following identity:

$$f_{0,\,n+i} + \sum_k f_{n+i,\,k}\, y'_k + \sum_k f_{n+i,\,n+k}\, y''_k = f_i. \tag{18}$$

We now multiply $(F)$ through by $y'_{1j}$, sum with respect to $j$, interchange the indices $j$ and $k$ on the left side and add to it equations (18) with $x_1$ substituted for $x$; in this way we obtain $(B)$. Using $y'_{2j}$ and $x_2$ in analogous manner enables us to show that $(H)$ implies $(E)$. It remains therefore to prove $(G)$, $(F)$, and $(H)$.

To prove the last two relations, we remember that $u_{jk}$, and also $v_{jk}$, are conjugate systems of solutions of the differential equations (4); hence we have†

* $(A)$ may also be derived from $(C)$.

† Bolza, p. 626; Hadamard, loc. cit., p. 319

$$\psi(u_j; u_i) = 0,$$

where $u_j$ and $u_i$ designate the systems $(u_{j1}, \cdots, u_{jn})$ and $(u_{i1}, \cdots, u_{in})$, and where

$$\psi(\eta; \zeta) = \sum_k \left[ \eta_k \frac{\partial \Omega}{\partial \zeta'_k} - \zeta_k \frac{\partial \Omega}{\partial \eta'_k} \right].$$

The equation $\psi(u_j; u_i) = 0$ reduces to

$$\sum_{kl} u_{jl}(Q_{kl} u_{ik} + R_{kl} u'_{ik}) - \sum_{kl} u_{il}(Q_{kl} u_{jk} + R_{kl} u'_{jk}) = 0.$$

If, in this relation we put $x$ equal to $x_1$ and make use of the initial conditions to which the functions $u_{jk}$ are subject, we obtain

$$Q_{ij} + \sum_k R_{kj} u'_{ik} = Q_{ji} + \sum_k R_{ki} u'_{jk},$$

which is equivalent to $(F)$ in virtue of the meaning of the symbols $Q_{ij}$ and $R_{ij}$. The corresponding treatment for the functions $v_{jk}$ will prove the relation $(H)$. Finally we have for any two systems of functions $u_{ik}$ and $v_{jk}$ the relation*

$$\sum_{kl} v_{jl}(Q_{kl} u_{ik} + R_{kl} u'_{ik}) - \sum_{kl} u_{il}(Q_{kl} v_{jk} + R_{kl} v'_{jk}) = \text{const.}$$

If we put $x$ successively equal to $x_1$ and to $x_2$ and set the two results equal to each other, making use at the same time of the initial conditions on the functions $u_{jk}$ and $v_{jk}$, we obtain $(G)$.

### 3. The case of one variable endpoint

We suppose now that the endpoint $P_1$ of our extremal $P_1 P_2$ is allowed to move along a curve $L$ in the space of $n + 1$ dimensions, whose equations are

$$x = \xi(t), \qquad y_i = \eta_i(t), \qquad t' \leqq t \leqq t''. \tag{19}$$

We suppose that this curve has no singular points, i. e.,

$$\xi'^2(t) + \sum_i [\eta'_i(t)]^2 \neq 0, \qquad t' \leqq t \leqq t'', \tag{20}$$

and that it is of class $C''$ in the interval $(t' t'')$.

We assume furthermore that we have found an extremal $E$

$$y_i = y_i(x), \qquad x_1 \leqq x \leqq x_2, \tag{21}$$

which satisfies the following conditions:

I. *The condition of Clebsch*:†

* Bolza, p. 626; Hadamard, loc. cit., p. 319.

† Bolza, p. 608.

At every point of the extremal the quadratic form $\sum_{ik} R_{ik}\,\zeta_i\,\zeta_k$ must be positive, i. e.:

$$\sum_{ik} R_{ik}\,\zeta_i\,\zeta_k > 0, \qquad x_1 \leqq x \leqq x_2. \tag{22}$$

II. *The condition of Mayer*:*

$$x_2 \leqq x_1',$$

where $x_1'$ is defined as the root next following upon $x_1$ of Mayer's determinant (14). The point $P_1'$ determined upon $E$ by $x_1'$ is known as the conjugate point.

III. The value $t_1$ of $t$ furnishes on $L$ the point where it intersects the extremal $E$; i. e.,

$$\xi_1 = \xi(t_1) = x_1, \qquad \eta_{1i} = \eta_i(t_1) = y_i(x_1) = y_{1i}.$$

IV. The function $f$ does not vanish at the point of intersection of the curves $L$ and $E$; i. e., $f^{(1)} = f(x_1; y_1; y_1') \neq 0$.

We ask now what further conditions the extremal $E$ must satisfy in order to furnish a minimum for the integral (1) among all admissible curves which join a point on $L$ to $P_2$. For this purpose we compute the extremal integral along $E$. Since the coördinates of $P_2$ are fixed and the coördinates of $P_1$ are all functions of the single variable $t$, the extremal integral will become a function of $t$ alone,

$$I(x_1; y_1; x_2; y_2) = I\big(\xi(t); \eta(t); x_2; y_2\big) = \Gamma(t),$$

which must satisfy the ordinary conditions for a minimum at $t_1$, namely,

$$\Gamma'(t_1) = 0, \qquad \Gamma''(t_1) \geqq 0.$$

By the use of formulæ (12), we find:

$$\Gamma'(t_1) = \frac{\partial I}{\partial x_1}\xi' + \sum_i \frac{\partial I}{\partial y_{1i}}\eta_i'|_1 = -\,\xi_1' f^{(1)} + \sum_i f^{(1)}_{n+i}(\xi_1' y_{1i}' - \eta_{1i}').$$

Its vanishing requires, as is well known, that $L$ must cut the extremal $E$ transversally at $P_1$. This condition of transversality, together with condition IV implies that

$$\sum_i c_i^2 \neq 0, \tag{23}$$

where $c_i = \xi_1\, y_{1i}' - \eta_{1i}'$.

A first value for $\Gamma''(t)$ is given by the formula

$$\Gamma''(t) = \frac{\partial I}{\partial x_1}\xi'' + \sum_i \frac{\partial I}{\partial y_{1i}}\eta_i'' + \frac{\partial^2 I}{\partial x_1^2}\xi'^2 + 2\sum_i \frac{\partial^2 I}{\partial x_1\,\partial y_{1i}}\xi'\,\eta_i' + \sum_{ij} \frac{\partial^2 I}{\partial y_{1i}\,\partial y_{1j}}\eta_i'\,\eta_j'. \tag{24}$$

* Ibid., pp. 610–619.

In the expressions (17) for the second derivatives of $I$ we replace the solutions $u_{kj}$ of the Jacobi equations by the fundamental system of solutions $\zeta_{rk}$ whose matrix reduces at $x = x_1$ to the identity matrix. Then it follows that

$$u_{jk} = \zeta_{jk} - \sum_{r=n+1}^{2n} \frac{Z^{(jk)}(x_2)}{Z(x_2)} \zeta_{rk}, \qquad u'_{jk}(x_1) = -\frac{Z^{(jk)}(x_2)}{Z(x_2)},$$

where $Z(x)$ is the determinant $|\zeta_{rk}|$ $(r = n+1, \cdots, 2n; k = 1, \cdots, n)$, and $Z^{(jk)}(x)$ is this determinant after its $k$th column has been replaced by $\zeta_{j1}, \cdots, \zeta_{jn}$. With the help of $(17_1)$, $(17_4)$, and $(17_8)$ equation (24) then takes the form

$$\Gamma''(t_1) = A(x_1) + B(x_1, x_2),$$

where

$$A(x_1) = -f^{(1)}\xi_1'' - f_0^{(1)}\xi_1'^2 + \sum_i f_{n+i}^{(1)}(\xi_1'' y'_{1i} - \eta''_{1i}) + \sum_i \xi_1' f_i^{(1)}(\xi_1' y'_{1i} - 2\eta'_{1i}) - \sum_{ij} f_{n+i,\,j}^{(1)} c_i c_j,$$

$$B(x_1, x_2) = \frac{1}{Z(x_2)} \sum_{ijk} Z^{(jk)}(x_2) R_{ik}(x_1) c_i c_j.$$

We investigate now the behavior of the function $B(x) = B(x_1, x)$ as $x$ varies from $x_1$ to $x_1'$, merely sketching in the briefest way the proofs of the results. We find

$$B'(x) = -\frac{1}{R(x_1)Z^2} \sum_{ij} \beta_{ij} c_i c_j,$$

where $R(x)$ is the determinant $|R_{ik}|$ and

$$\beta_{ij} = -R(x_1) \sum_k R_{ik}(x_1)[ZZ^{(jk)\prime} - Z' Z^{(jk)}] = \sum_{klts} R_{ik}(x_1) Z_{kl} \rho_{ls} Z_{ts} R_{tj}(x_1), \tag{25}$$

in which $Z_{kl}$ is the cofactor of $\zeta_{kl}$ in $Z(x)$, and $\rho_{ls}$ the cofactor of $R_{ls}(x)$ in $R(x)$. The last expression for $\beta_{ij}$ is derived from the first by applying successively formulæ (70), (65), and (55) in Chapter XII of Bolza's *Vorlesungen*, and remembering that the function $\psi$ of formula (55) is a constant which can be evaluated at $x_1$, where the values of the functions $\zeta_{rk}$ are prescribed. In these formulæ, the solutions $z$, $u$ of Jacobi's equations are to be replaced by solutions of the matrix $\zeta_{rk}$ defined above; and for our case the multipliers $\rho$, $r$ are all equal to zero identically, since there are no adjoined differential equations.

The quadratic form $\sum \beta_{ij} c_i c_j$ is positive, on account of condition (23) together with the fact that it is derived from the quadratic form $\sum \rho_{ls} x_l x_s$ by two linear transformations whose determinants are equal to $Z$ and $R$ respectively. We have $Z \neq 0$ because the Mayer condition is fulfilled and

because the Mayer determinant (14) reduces to $Z$ when the functions (5) are expressed linearly in terms of the functions $\zeta_{rk}$. Further $R \neq 0$ on account of condition I, and since the matrix $|\rho_{ls}|$ is the matrix of the cofactors of $R$, the form $\sum \rho_{ls} x_l x_s$ is positive. If we now apply condition I once more, we conclude that $B'(x)$ is steadily negative on the interval $x_1 < x < x_1'$.

The initial values of the functions $\zeta_{rk}$ at $x = x_1$ are known, and hence the first few terms of the Taylor expansions for the numerator and denominator of $B(x)$ can be found. It is found that in the neighborhood of $x_1$, we may write

$$B(x) = \frac{1}{x - x_1}\left\{\sum_{ij} R_{ij}(x_1)\, c_i\, c_j + a_1(x - x_1) + \cdots\right\},$$

so that $B(x)$ approaches $+\infty$ as $x$ approaches $x_1$ from the right.

Finally, setting $x$ equal to $x_1'$ in the two expressions for $\beta_{ij}$ given in formula (25), we obtain, since $Z(x_1') = 0$,

$$R(x_1)\sum_k R_{ik}(x_1)\, Z'\, Z^{(jk)} = \sum_{klls} R_{ik}(x_1)\, Z_{kl}\, \rho_{ls}\, Z_{ts}\, R_{tj}(x_1),$$

the argument $x$ being everywhere replaced by $x_1'$. It follows from this that the numerator of $B(x_1')$ reduces to

$$\frac{1}{R(x_1)\, Z'}\sum_{ij}\beta_{ij}\, c_i\, c_j,$$

which is different from zero; further, since the denominator vanishes, $B(x)$ becomes infinite as $x$ approaches $x_1'$, and it must approach $-\infty$, because $B'(x)$ is negative.

From these facts it follows that the function $B(x_1, x)$ decreases univariantly from $+\infty$ to $-\infty$ as $x$ increases from $x_1$ to $x_1'$. Hence there is an unique value $x_1''$ between $x_1$ and $x_1'$ such that $\Gamma''(t_1)$ is positive for every value $x_2$ satisfying the condition $x_1 < x_2 < x_1''$. The point defined on $E$ by $x_1''$ is called the "focal point" of $L$ on $E$. We have therefore the following theorem:

THEOREM II. *If an extremal arc $E$ with the properties I–IV minimizes the integral* (1) *with respect to arcs joining a fixed curve $L$ to a fixed point $P_2$, then it must be cut transversally by $L$ at their point of intersection $P_1$, and it must further be such that the focal point of $L$ on $E$ is not between $P_1$ and $P_2$.*

THE UNIVERSITY OF WISCONSIN
*September*, 1915

# DEFORMABLE TRANSFORMATIONS OF RIBAUCOUR*

BY

LUTHER PFAHLER EISENHART

When a system of spheres involves two parameters, their envelope consists in general of two sheets, say $\Sigma$ and $\Sigma_1$, and the centers of the spheres lie upon a surface $S$. A correspondence between $\Sigma$ and $\Sigma_1$ is established by the points of contact on the same sphere. In general the lines of curvature on $\Sigma$ and $\Sigma_1$ do not correspond. When they do, we say that $\Sigma$ is in the relation of a transformation of Ribaucour with $\Sigma_1$, and vice versa. For the sake of brevity we call it a *transformation* $R$.

It is a known property of envelopes of spheres that if $S$ be deformed and the spheres be carried along in the deformation, the points of contact of the spheres with their envelope in the new position are the same as before the deformation. Ordinarily when $S$ for a transformation $R$ is deformed, the sheets $\Sigma'$ and $\Sigma_1'$ of the new envelope are not in the relation of a transformation $R$. Bianchi† has shown that when $S$ is a surface applicable to a surface of revolution, it is possible to choose spheres so that for every deformation of $S$ the two sheets of the envelope shall be in the relation of a transformation $R$. From the equations which Bianchi used to prove this result it can be shown that if $S$ undergoes a single deformation so that the sheets of the new envelope are in the relation of a transformation $R$ the conjugate system on the deform of $S$ corresponding to the lines of curvature on the envelope is the same as before deformation. It is known that a conjugate system may be preserved in an infinity of deformations, or one, or none, the latter being the general case. It is the purpose of this paper to determine the transformations $R$ whose surfaces of center admit one deformation into surfaces of center of transformations $R$.

It is shown that the only surfaces $\Sigma$ admitting such transformations $R$ have the same spherical representation of their lines of curvature as isothermic surfaces, and that every surface of this type admits such transformations.

The conjugate system on the surface $S$ is a system $2O$, to use the notation of Guichard. A conjugate system remaining conjugate in a single deformation is called a *permanent conjugate system*. The present investigation carries

* Presented to the Society, at Cambridge, Sept. 4, 1916.

† *Differenziale geometria*, vol. 2, p. 117.

with it the determination of all permanent conjugate systems which are $2O$. It shows also that any conjugate system having the same spherical representation as a permanent system $2O$ is of this kind also.

Evidently when one sheet of an envelope of spheres is a point or a plane the transformation is of the type $R$. Hence a special class of the solutions of our problem include all transformations $R$ whose surfaces of center may be deformed so that in the new position the spheres pass through a point or are tangent to a plane, which we have shown to be respectively transformations $D_m$ of isothermic surfaces,* and transformations $E_m$ of surfaces with isothermal representations of their lines of curvature.†

When our general results are applied to the transformations $D_m$ of isothermic surfaces, we are led to the transformations $T_m$ of these surfaces discovered by Bianchi‡ and expressed intrinsically. The present results give a geometrical basis to these transformations and other observations made recently by Bianchi.§

## 1. Transformations of Ribaucour

Let $\Sigma$ be a surface referred to its lines of curvature, $u = \text{const.}$, $v = \text{const.}$; $x, y, z$, its cartesian coördinates; $X, Y, Z$, the direction-cosines of its normal; $\rho_1, \rho_2$, its principal radii of normal curvature. These quantities satisfy the following equations of Rodrigues:‖

$$\frac{\partial x}{\partial u} + \rho_1 \frac{\partial X}{\partial u} = 0, \qquad \frac{\partial x}{\partial v} + \rho_2 \frac{\partial X}{\partial v} = 0. \tag{1}$$

Darboux¶ has shown that if $\lambda$ and $\mu$ are any two functions satisfying the equations

$$\frac{\partial \lambda}{\partial u} + \rho_1 \frac{\partial \mu}{\partial u} = 0, \qquad \frac{\partial \lambda}{\partial v} + \rho_2 \frac{\partial \mu}{\partial v} = 0, \tag{2}$$

on the surface $S$ whose coördinates $x_0, y_0, z_0$ are given by

$$x_0 = x - \frac{\lambda}{\mu} X, \qquad y_0 = y - \frac{\lambda}{\mu} Y, \qquad z_0 = z - \frac{\lambda}{\mu} Z, \tag{3}$$

the parametric curves form a conjugate system, and that the spheres with centers on $S$ and whose radii are given by $\lambda/\mu$ give a transformation $R$ of $\Sigma$, and the most general one. We shall now find the expressions for the cartesian coördinates $x_1, y_1, z_1$ of the transform $\Sigma_1$.

* Rendiconti dei Lincei, ser. 5, vol. 24 (1915), p. 351.
† Annals of Mathematics, ser. 2, vol. 17 (1915), pp. 64–71.
‡ Annali di Matematica, ser. 3, vol. 12 (1906), pp. 19–54.
§ Rendiconti dei Lincei, ser. 5, vol. 24 (1915), pp. 377–387.
‖ E., p. 122. A reference of this sort is to the author's *Differential Geometry*.
¶ *Lecons*, vol. 2, p. 383.

If $X', Y', Z'; X'', Y'', Z''$ denote the direction-cosines of the tangents to the curves $v = \text{const.}$, $u = \text{const.}$ respectively on $\Sigma$, we have*

$$X' = \frac{1}{\sqrt{E}}\frac{\partial x}{\partial u}, \qquad X'' = \frac{1}{\sqrt{G}}\frac{\partial x}{\partial v},$$

$$\frac{\partial X'}{\partial u} = -\frac{1}{\sqrt{G}}\frac{\partial \sqrt{E}}{\partial v}X'' + \frac{\sqrt{E}}{\rho_1}X, \qquad \frac{\partial X'}{\partial v} = \frac{1}{\sqrt{E}}\frac{\partial \sqrt{G}}{\partial u}X'', \tag{4}$$

$$\frac{\partial X''}{\partial u} = \frac{1}{\sqrt{G}}\frac{\partial \sqrt{E}}{\partial v}X', \qquad \frac{\partial X''}{\partial v} = -\frac{1}{\sqrt{E}}\frac{\partial \sqrt{G}}{\partial u}X' + \frac{\sqrt{G}}{\rho_2}X,$$

$$\frac{\partial X}{\partial u} = -\frac{\sqrt{E}}{\rho_1}X', \qquad \frac{\partial X}{\partial v} = -\frac{\sqrt{G}}{\rho_2}X''.\dagger$$

The equation of the sphere of radius $\lambda/\mu$ and center (3) is reducible to

$$\mu \sum (x_1 - x)^2 + 2\lambda \sum X(x_1 - x) = 0,$$

where $x_1, y_1, z_1$ are current coördinates. The equations obtained by differentiating this equation with respect to $u$ and $v$ are

$$\frac{\partial \lambda}{\partial u}\sum (x_1 - x)^2 + 2\lambda\sqrt{E}\sum X'(x_1 - x) = 0,$$

$$\frac{\partial \lambda}{\partial v}\sum (x_1 - x)^2 + 2\lambda\sqrt{G}\sum X''(x_1 - x) = 0.$$

These three equations may be replaced by

$$x_1 - x = -\frac{1}{\sigma m}(\alpha X' + \beta X'' + \mu X), \tag{5}$$

where $m$ is a constant, and the functions $\alpha, \beta, \sigma$ are subject to the conditions

$$\alpha^2 + \beta^2 + \mu^2 = 2m\lambda\sigma, \tag{6}$$

and

$$\frac{\partial \lambda}{\partial u} = \alpha\sqrt{E}, \qquad \frac{\partial \lambda}{\partial v} = \beta\sqrt{G}.$$

Comparing equations (1) and (2), we see that $\lambda$ is a solution of the equation satisfied by $x, y, z$, that is, the point equation for $\Sigma$. Expressing this condition, we find the expressions for $\partial\alpha/\partial v$ and $\partial\beta/\partial u$ in the fundamental equations (I) below. The third and fourth of equations (I) below follow from the

* E., p. 157.

† Algebraic signs are given to these direction-cosines so that we have

$$\begin{vmatrix} X' & Y' & Z' \\ X'' & Y'' & Z'' \\ X & Y & Z \end{vmatrix} = 1.$$

first two and (2). The functions $k$ and $l$ in the fifth and sixth of equations (I) may be taken as defined by these equations. The expressions for $\partial\alpha/\partial u$ and $\partial\beta/\partial v$ are a consequence of the differentiation of (6). The fundamental system is

$$
\begin{aligned}
&\frac{\partial\lambda}{\partial u} = \sqrt{E}\alpha, \qquad \frac{\partial\lambda}{\partial v} = \sqrt{G}\beta,\\
&\frac{\partial\mu}{\partial u} = -\frac{\sqrt{E}}{\rho_1}\alpha, \qquad \frac{\partial\mu}{\partial v} = -\frac{\sqrt{G}}{\rho_2}\beta,\\
&\frac{\partial \log\sigma}{\partial u} = k\frac{\alpha}{\lambda}, \qquad \frac{\partial \log\sigma}{\partial v} = l\frac{\beta}{\lambda},\\
&\frac{\partial\alpha}{\partial u} = -\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v}\beta + \frac{\sqrt{E}}{\rho_1}\mu + m(k+\sqrt{E})\sigma,\\
&\frac{\partial\alpha}{\partial v} = \frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u}\beta, \qquad \frac{\partial\beta}{\partial u} = \frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v}\alpha,\\
&\frac{\partial\beta}{\partial v} = -\frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u}\alpha + \frac{\sqrt{G}}{\rho_2}\mu + m(l+\sqrt{G})\sigma,\\
&\frac{\partial k}{\partial v} = l\left[\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v} - \frac{\beta}{\lambda}(k+\sqrt{E})\right],\\
&\frac{\partial l}{\partial u} = k\left[\frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u} - \frac{\alpha}{\lambda}(l+\sqrt{G})\right].
\end{aligned}
\tag{I}
$$

The last two of these equations must be satisfied in order that the conditions of integrability of the derivatives of $\alpha$ and $\beta$ shall be satisfied; in deriving them we make use of the Gauss and Codazzi equations for $\Sigma$, namely

$$
\begin{aligned}
&\frac{\partial}{\partial v}\left(\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v}\right) + \frac{\partial}{\partial u}\left(\frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u}\right) + \frac{\sqrt{EG}}{\rho_1\rho_2} = 0,\\
&\frac{\partial}{\partial v}\left(\frac{\sqrt{E}}{\rho_1}\right) = \frac{1}{\rho_2}\frac{\partial\sqrt{E}}{\partial v}, \qquad \frac{\partial}{\partial u}\left(\frac{\sqrt{G}}{\rho_2}\right) = \frac{1}{\rho_1}\frac{\partial\sqrt{G}}{\partial u}.
\end{aligned}
\tag{7}
$$

Equations (I) form a completely integrable system of differential equations and each set of solutions satisfying (6) gives a surface $\Sigma_1$, which, as we shall show, is a Ribaucour transform of $\Sigma$.

From (5) we have by differentiation

$$
\frac{\partial x_1}{\partial u} = -kX_1', \qquad \frac{\partial x_1}{\partial v} = lX_1'', \tag{8}
$$

where

$$X_1' = \left(1 - \frac{\alpha^2}{m\sigma\lambda}\right) X' - \frac{\alpha\beta}{m\sigma\lambda} X'' - \frac{\alpha\mu}{m\sigma\lambda} X,$$

(9)

$$X_1'' = \frac{\alpha\beta}{m\sigma\lambda} X' - \left(1 - \frac{\beta^2}{m\sigma\lambda}\right) X'' + \frac{\beta\mu}{m\sigma\lambda} X.$$

Similar equations hold for the $y$'s and $z$'s. Now $X_1'$, $Y_1'$, $Z_1'$; $X_1''$, $Y_1''$, $Z_1''$ are the direction-cosines of the tangents to the curves $v$ = const., $u$ = const. respectively on $\Sigma_1$. From these we find that the direction-cosines, $X_1$, $Y_1$, $Z_1$, of the normal to $\Sigma_1$ are of the form

$$X_1 = -\frac{\alpha\mu}{m\sigma\lambda} X' - \frac{\beta\mu}{m\sigma\lambda} X'' + \left(1 - \frac{\mu^2}{m\sigma\lambda}\right) X. \tag{10}$$

If $E_1$ and $G_1$ are the first fundamental coefficients of $\Sigma_1$, we have on comparing (8) with (4) that

$$\sqrt{E_1} = -k, \qquad \sqrt{G_1} = l. \tag{11}$$

On differentiating (10) we get

$$\frac{\partial X_1}{\partial u} + \frac{1}{\rho_{11}} \frac{\partial x_1}{\partial u} = 0, \qquad \frac{\partial X_1}{\partial v} + \frac{1}{\rho_{12}} \frac{\partial x_1}{\partial v} = 0,$$

where $\rho_{11}$ and $\rho_{12}$ are given by

$$k\left(\frac{1}{\rho_{11}} + \frac{\mu}{\lambda}\right) + \sqrt{E}\left(\frac{1}{\rho_1} + \frac{\mu}{\lambda}\right) = 0,$$

(12)

$$l\left(\frac{1}{\rho_{12}} + \frac{\mu}{\lambda}\right) + \sqrt{G}\left(\frac{1}{\rho_2} + \frac{\mu}{\lambda}\right) = 0.$$

Since the preceding equations are analogous to (1), we know that the lines of curvature on $\Sigma_1$ are parametric, and that $\rho_{11}$ and $\rho_{12}$, given by (12), are the principal radii of normal curvature of $\Sigma_1$. Hence

THEOREM 1. *Every set of solutions of equations* (I) *which satisfy the quadratic relation* (6) *determines a transformation* $R$ *of* $\Sigma$.

In consequence of (10) equation (5) can be given the form

$$x_1 - x = \frac{\mu}{\lambda}(X_1 - X). \tag{13}$$

## 2. THE TRANSFORMATION $R$ OBTAINED BY THE DEFORMATION OF $S$

From (I) we find

$$\frac{\partial}{\partial u}\left(\frac{\lambda}{\mu}\right) = \frac{\alpha}{\mu} L_1, \qquad \frac{\partial}{\partial v}\left(\frac{\lambda}{\mu}\right) = \frac{\beta}{\mu} L_2, \tag{14}$$

where we have put

$$L_1 = \sqrt{E}\left(1 + \frac{1}{\rho_1}\frac{\lambda}{\mu}\right), \qquad L_2 = \sqrt{G}\left(1 + \frac{1}{\rho_2}\frac{\lambda}{\mu}\right). \tag{15}$$

Hence the first fundamental coefficients of $S$, whose coördinates are given by (3), have the form

$$E_0 = L_1^2\left(1 + \left(\frac{\alpha}{\mu}\right)^2\right), \qquad F_0 = L_1 L_2 \frac{\alpha}{\mu}\frac{\beta}{\mu}, \qquad G_0 = L_2^2\left(1 + \left(\frac{\beta}{\mu}\right)^2\right). \tag{16}$$

When $S$ is deformed into the surface of centers $S'$ of another transformation $R$, the functions of the latter must necessarily give equivalent expressions for $E_0$, $F_0$, and $G_0$. Since the spheres are unaltered in size, the functions $\lambda'$ and $\mu'$ are given by

$$\lambda' = \rho\lambda, \qquad \mu' = \rho\mu, \tag{17}$$

where $\rho$ is a factor of proportionality.*

We call $\Sigma'$ and $\Sigma_1'$ the two sheets of the new transformation $R$ and indicate by primes all functions belonging to the latter. From (14) and analogous equations we have

$$\begin{aligned} \sqrt{E}\frac{\alpha}{\mu}\left(1 + \frac{1}{\rho_1}\frac{\lambda}{\mu}\right) &= \sqrt{E'}\frac{\alpha'}{\mu'}\left(1 + \frac{1}{\rho_1'}\frac{\lambda}{\mu}\right), \\ \sqrt{G}\frac{\beta}{\mu}\left(1 + \frac{1}{\rho_2}\frac{\lambda}{\mu}\right) &= \sqrt{G'}\frac{\beta'}{\mu'}\left(1 + \frac{1}{\rho_2'}\frac{\lambda}{\mu}\right). \end{aligned} \tag{18}$$

In order that the expressions (16) may be equal to similar ones for $S'$, we must have in consequence of (18)

$$\frac{\alpha'^2}{\mu'^2} = \frac{\alpha^2}{\mu^2}, \qquad \frac{\beta'^2}{\mu'^2} = \frac{\beta^2}{\mu^2}.$$

Because of (17) these may be replaced by

$$\alpha' = \rho\alpha, \qquad \beta' = -\rho\beta.\dagger \tag{19}$$

When these values are substituted in equations analogous to the first two of (I), we find for the determination of $\rho$

$$\frac{\partial \log \rho}{\partial u} = \frac{\alpha}{\lambda}(\sqrt{E'} - \sqrt{E}), \qquad \frac{\partial \log \rho}{\partial v} = -\frac{\beta}{\lambda}(\sqrt{G'} + \sqrt{G}). \tag{20}$$

* By thus fixing the signs of $\lambda'$ and $\mu'$ we have merely chosen the positive direction of the normal to $\Sigma'$.

† The choice of these signs is merely equivalent to determining the signs of the square roots of $E'$ and $G'$ hereafter.

Moreover, equations (18) are reducible to

$$\begin{aligned}\sqrt{E'}\left(1+\frac{1}{\rho_1}\frac{\lambda}{\mu}\right)&=\sqrt{E}\left(1+\frac{1}{\rho_1}\frac{\lambda}{\mu}\right),\\ -\sqrt{G'}\left(1+\frac{1}{\rho_2}\frac{\lambda}{\mu}\right)&=\sqrt{G}\left(1+\frac{1}{\rho_2}\frac{\lambda}{\mu}\right).\end{aligned}\tag{21}$$

If these equations be differentiated with respect to $v$ and $u$ respectively and in the reduction use be made of the Codazzi equations (7) and similar ones for $\Sigma'$, we get

$$\begin{aligned}\frac{1}{\sqrt{G'}}\frac{\partial\sqrt{E'}}{\partial v}+\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v}&=-\frac{\beta}{\lambda}(\sqrt{E'}-\sqrt{E}),\\ \frac{1}{\sqrt{E'}}\frac{\partial\sqrt{G'}}{\partial u}+\frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u}&=\frac{\alpha}{\lambda}(\sqrt{G'}+\sqrt{G}).\end{aligned}\tag{22}$$

When these expressions are substituted in the Gauss equation for $\Sigma'$, analogous to (7), we are led to the condition

$$2\sqrt{E'G'}-\sqrt{E'}(l-\sqrt{G})+\sqrt{G'}(k-\sqrt{E})+\sqrt{E}l+\sqrt{G}k=0.\tag{23}$$

If we choose the constant $m'$ of the new transformation $R$ equal to $m$, it follows from the foregoing results and (6) that

$$\sigma'=\rho\sigma.\tag{24}$$

In order that $\mu'$, $\alpha'$, $\beta'$, and $\sigma'$, as given by (17), (19), and (24), shall satisfy equations for $\Sigma'$ analogous to (I) we must have

$$\begin{aligned}k'-k+\sqrt{E}-\sqrt{E'}&=0,\\ l'+l-\sqrt{G}-\sqrt{G'}&=0.\end{aligned}\tag{25}$$

Moreover, these equations satisfy equations similar to the last two of (I) when (23) is satisfied.

### 3. Equations in reduced form. The inverse of a transformation $R$

Equation (23) may be replaced by the two equations

$$\begin{aligned}\sqrt{E'}&=\frac{1}{2}(\sqrt{E}-k)+\frac{e^{\theta}}{2}(\sqrt{E}+k),\\ -\sqrt{G'}&=\frac{1}{2}(\sqrt{G}-l)+\frac{e^{-\theta}}{2}(\sqrt{G}+l),\end{aligned}\tag{26}$$

where $\theta$ is a function to be determined by the equation (22). Reviewing the

preceding discussion, we note that when such a function $\theta$ is known, the surface $\Sigma'$ is defined intrinsically by (26) and (21).

When the values (26) of $\sqrt{E'}$ and $\sqrt{G'}$ are substituted in (22), we get for the determination of $\theta$

$$(\sqrt{E}+k)e^{\theta}\frac{\partial\theta}{\partial v}=(e^{\theta}-e^{-\theta})(\sqrt{G}+l)\left(\frac{\beta}{2\lambda}(\sqrt{E}+k)-\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v}\right),$$

(27)

$$(\sqrt{G}+l)e^{-\theta}\frac{\partial\theta}{\partial u}=(e^{\theta}-e^{-\theta})(\sqrt{E}+k)\left(\frac{\alpha}{2\lambda}(\sqrt{G}+l)-\frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u}\right).$$

These equations are satisfied by $e^{\theta}=1$. In this case $\Sigma'$ is congruent to $\Sigma$. Another constant solution is $e^{\theta}=-1$. Now equations (26) become

$$\sqrt{E'}=-k, \qquad \sqrt{G'}=l. \tag{28}$$

Comparing these equations with (11) and (21) with (12), we see that $\Sigma'$ is $\Sigma_1$, the transform of $\Sigma$, and consequently its transform $\Sigma_1'$ is $\Sigma$. By means of the foregoing results we shall then be able to find the expressions for the functions $\lambda^{-1}$, $\mu^{-1}$, $\alpha^{-1}$, $\beta^{-1}$, $\sigma^{-1}$ by which $\Sigma$ is obtained from $\Sigma_1$. In fact, from (20) we have, to within a negligible constant factor, $\rho=1/\lambda\sigma$. Consequently from (17), (19), and (24) we have

$$\lambda^{-1}=\frac{1}{\sigma}, \qquad \mu^{-1}=\frac{\mu}{\lambda\sigma}, \qquad \alpha^{-1}=\frac{\alpha}{\lambda\sigma}, \qquad \beta^{-1}=-\frac{\beta}{\lambda\sigma}, \qquad \sigma^{-1}=\frac{1}{\lambda}. \tag{29}$$

It is readily shown that these expressions satisfy equations for $\Sigma_1$ analogous to (I).

## 4. Solution of equations (27)

In view of the foregoing results, we are concerned with the solutions of equations (27) other than $e^{2\theta}=1$. By means of (I) these equations are reducible to

$$\frac{\partial\theta}{\partial u}=(1-e^{2\theta})\frac{\partial B}{\partial u}, \qquad \frac{\partial\theta}{\partial v}=-(1-e^{-2\theta})\frac{\partial A}{\partial v}, \tag{30}$$

where

$$A=\log\sqrt{\frac{\sigma}{\lambda}}(\sqrt{E}+k), \qquad B=\log\sqrt{\frac{\sigma}{\lambda}}(\sqrt{G}+l). \tag{31}$$

The condition of integrability of (30) is reducible to

$$e^{2\theta}L-M=0, \tag{32}$$

where

$$L=\frac{\partial^2 B}{\partial u\partial v}-2\frac{\partial A}{\partial v}\frac{\partial B}{\partial u}, \qquad M=\frac{\partial^2 A}{\partial u\partial v}-2\frac{\partial A}{\partial v}\frac{\partial B}{\partial u}. \tag{33}$$

If $L=M=0$, there are an infinity of deformations of the kind sought.

Bianchi has solved this problem and hence we exclude it from our consideration.

Substituting the expression for $e^{2\theta}$ from (32) in (30), we get

$$\frac{\partial}{\partial u}\left(\frac{L}{M}\right) = 2\left(1 - \frac{L}{M}\right)\frac{\partial B}{\partial u}, \qquad \frac{\partial}{\partial v}\left(\frac{M}{L}\right) = 2\left(1 - \frac{M}{L}\right)\frac{\partial A}{\partial v}.$$

The integrals of these equations are

$$\frac{L}{M} = 1 - Ve^{-2B}, \qquad \frac{M}{L} = 1 - Ue^{-2A},$$

where $U$ and $V$ are functions of $u$ and $v$ respectively.

Since we have excluded the case $e^{2\theta} = 1$, neither $U$ nor $V$ can be zero. The consistency of these equations necessitates

$$UV - Ue^{2B} - Ve^{2A} = 0,$$

which in consequence of (31) is equivalent to

$$\frac{\lambda}{\sigma}UV - (\sqrt{G} + l)^2 U - (\sqrt{E} + k)^2 V = 0.$$

Because of the significance of $k$ and $l$ as shown by (11), we know that the parameters $u$ and $v$ of the lines of curvature on $\Sigma$ can be chosen so that $U$ and $V$ may be replaced by constants, which it is convenient to take as $\pm 4$. Hence we have for consideration the three cases

$$(\sqrt{E} + k)^2 + (\sqrt{G} + l)^2 = \pm 4\frac{\lambda}{\sigma}, \qquad U = V = \pm 4, \tag{34}$$

$$(\sqrt{E} + k)^2 - (\sqrt{G} + l)^2 = 4\frac{\lambda}{\sigma}, \qquad U = -V = 4, \tag{35}$$

$$(\sqrt{E} + k)^2 - (\sqrt{G} + l)^2 = -4\frac{\lambda}{\sigma}, \qquad U = -V = -4. \tag{36}$$

If we restrict our consideration to the case of real transformations of real surfaces, we may replace (34) by

$$(\sqrt{E} + k) = 2\sqrt{\frac{\lambda\epsilon}{\sigma}}\cos\omega, \qquad (\sqrt{G} + l) = 2\sqrt{\frac{\lambda\epsilon}{\sigma}}\sin\omega, \tag{37}$$

where $\omega$ is a function to be determined and $\epsilon = \pm 1$ according as $\lambda/\sigma$ is positive or negative. Now (31) become

$$A = \log 2\sqrt{\epsilon}\cos\omega, \qquad B = \log 2\sqrt{\epsilon}\sin\omega,$$

and from (32) we get

$$e^{2\theta} = -\tan^2\omega.$$

These values of $A$ and $B$ satisfy (32) when $L$ and $M$ are given the values (33). Since $\theta$ is necessarily imaginary, equation (34) does not give rise to a real solution of the problem.

Before considering equation (36), we remark that if we put

$$\bar{\alpha} = \alpha, \quad \bar{\beta} = \beta, \quad \bar{\mu} = \mu, \quad \bar{\lambda} = \lambda, \quad \bar{\sigma} = -\sigma, \quad \bar{m} = -m,$$

equations (I) and (6) are satisfied, and from (5) it follows that the transform $\bar{\Sigma}_1$ is the same as $\Sigma_1$. Hence equation (36) offers nothing different from (35), which we now proceed to investigate.

If we replace (35) by

$$\sqrt{E} + k = 2\sqrt{\frac{\lambda\epsilon}{\sigma}}\cosh\omega, \qquad \sqrt{G} + l = 2\sqrt{\frac{\lambda\epsilon}{\sigma}}\sinh\omega, \tag{38}$$

where $\epsilon = \pm 1$ according as $\lambda/\sigma$ is positive or negative, we have

$$A = \log 2\sqrt{\epsilon}\cosh\omega, \qquad B = \log 2\sqrt{\epsilon}\sinh\omega,$$

and

$$e^{2\theta} = \tanh^2\omega. \tag{39}$$

From (I) we obtain

$$\frac{\partial \log\lambda\sigma}{\partial u} = (\sqrt{E} + k)\frac{\alpha}{\lambda}, \qquad \frac{\partial \log\lambda\sigma}{\partial v} = (\sqrt{G} + l)\frac{\beta}{\lambda}, \tag{40}$$

and

$$\begin{aligned}
\frac{\partial}{\partial u}\log\sqrt{\frac{\sigma}{\lambda}}(\sqrt{G} + l) &= \frac{\sqrt{E} + k}{\sqrt{G} + l}\frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u} - \frac{1}{2}(\sqrt{E} + k)\frac{\alpha}{\lambda}, \\
\frac{\partial}{\partial v}\log\sqrt{\frac{\sigma}{\lambda}}(\sqrt{E} + k) &= \frac{\sqrt{G} + l}{\sqrt{E} + k}\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v} - \frac{1}{2}(\sqrt{G} + l)\frac{\beta}{\lambda}.
\end{aligned} \tag{41}$$

Substituting the expressions from (38), we get

$$\begin{aligned}
\frac{\partial\omega}{\partial u} &= \frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u} - \frac{\alpha}{\sqrt{\lambda\sigma\epsilon}}\sinh\omega, \\
\frac{\partial\omega}{\partial v} &= \frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v} - \frac{\beta}{\sqrt{\lambda\sigma\epsilon}}\cosh\omega.
\end{aligned} \tag{42}$$

By means of (I) and (40) we find that the condition of integrability of these equations reduces to

$$\frac{\partial}{\partial u}\left(\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v}\right) = \frac{\partial}{\partial v}\left(\frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u}\right). \tag{43}$$

We shall now determine under what conditions a surface satisfies (43).

## 5. Surfaces with the same spherical representation of their lines of curvature as isothermic surfaces

If the linear element of the spherical representation of $\Sigma$ be written

$$d\sigma^2 = \mathcal{E}du^2 + \mathcal{G}dv^2,$$

then

$$\mathcal{E} = \rho_1^2 E, \qquad \mathcal{G} = \rho_2^2 G, \tag{44}$$

as follows from (1). Hence the Codazzi equations (7) may be written

$$\frac{1}{\sqrt{\mathcal{G}}}\frac{\partial\sqrt{\mathcal{E}}}{\partial v} = \frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v}, \qquad \frac{1}{\sqrt{\mathcal{E}}}\frac{\partial\sqrt{\mathcal{G}}}{\partial u} = \frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u},$$

and consequently equation (43) may be replaced by

$$\frac{\partial}{\partial u}\left(\frac{1}{\sqrt{\mathcal{G}}}\frac{\partial\sqrt{\mathcal{E}}}{\partial v}\right) = \frac{\partial}{\partial v}\left(\frac{1}{\sqrt{\mathcal{E}}}\frac{\partial\sqrt{\mathcal{G}}}{\partial u}\right). \tag{45}$$

It is our purpose to show that an orthogonal system on the unit sphere satisfying this condition represents the lines of curvature of two isothermic surfaces, which are the *Christoffel transforms* of one another by definition.

In the case of an isothermic surface, equation (44) may be replaced by

$$\rho_1 = \frac{a}{\sqrt{\mathcal{E}}}, \qquad \rho_2 = \frac{a}{\sqrt{\mathcal{G}}}, \tag{46}$$

or

$$\rho_1 = \frac{b}{\sqrt{\mathcal{E}}}, \qquad \rho_2 = -\frac{b}{\sqrt{\mathcal{G}}}, \tag{47}$$

where

$$E = G = a^2, \quad \text{or} \quad b^2. \tag{48}$$

The Codazzi equations (7) may be given the form

$$\frac{\partial \log\sqrt{\mathcal{E}}}{\partial v} = \frac{1}{\rho_2 - \rho_1}\frac{\partial\rho_1}{\partial v}, \qquad \frac{\partial \log\sqrt{\mathcal{G}}}{\partial u} = \frac{1}{\rho_1 - \rho_2}\frac{\partial\rho_2}{\partial u}.$$

Substituting in these equations from (46) and (47), we have respectively

$$\begin{aligned} &\frac{\partial \log a}{\partial v} = \frac{1}{\sqrt{\mathcal{G}}}\frac{\partial\sqrt{\mathcal{E}}}{\partial v}, \qquad \frac{\partial \log a}{\partial u} = \frac{1}{\sqrt{\mathcal{E}}}\frac{\partial\sqrt{\mathcal{G}}}{\partial u}, \\ &\frac{\partial \log b}{\partial v} = -\frac{1}{\sqrt{\mathcal{G}}}\frac{\partial\sqrt{\mathcal{E}}}{\partial v}, \qquad \frac{\partial \log b}{\partial u} = -\frac{1}{\sqrt{\mathcal{E}}}\frac{\partial\sqrt{\mathcal{G}}}{\partial u}. \end{aligned} \tag{49}$$

Equation (45) expresses the necessary and sufficient condition that these equations shall be consistent. Conversely, when the spherical representation of the lines of curvature of a surface satisfies (45), one obtains by quadratures

from (48) and (49) the first fundamental coefficients of two isothermic surfaces with the same spherical representation of their lines of curvature as the given surface. Hence we have

THEOREM 2. *The necessary and sufficient condition that a surface be isothermic or have the same spherical representation of its lines of curvature as an isothermic surface is that condition* (43) *be satisfied.*

## 6. TRANSFORMATIONS $R$ OF THE TYPE SOUGHT

Let $\Sigma$ be a surface satisfying equation (43). Then equations (I) may be replaced by

$$
\begin{aligned}
&\frac{\partial \lambda}{\partial u} = \sqrt{E}\alpha, \qquad \frac{\partial \lambda}{\partial v} = \sqrt{G}\beta,\\
&\frac{\partial \mu}{\partial u} = -\frac{\sqrt{E}}{\rho_1}\alpha, \qquad \frac{\partial \mu}{\partial v} = -\frac{\sqrt{G}}{\rho_2}\beta,\\
&\frac{\partial \alpha}{\partial u} = -\frac{1}{\sqrt{G}}\frac{\partial \sqrt{E}}{\partial v}\beta + \mu\frac{\sqrt{E}}{\rho_1} + 2m\sqrt{\lambda\sigma\epsilon}\cosh\omega,\\
&\frac{\partial \alpha}{\partial v} = \frac{\beta}{\sqrt{E}}\frac{\partial \sqrt{G}}{\partial u}, \qquad \frac{\partial \beta}{\partial u} = \frac{\alpha}{\sqrt{G}}\frac{\partial \sqrt{E}}{\partial v},\\
&\frac{\partial \beta}{\partial v} = -\frac{1}{\sqrt{E}}\frac{\partial \sqrt{G}}{\partial u}\alpha + \mu\frac{\sqrt{G}}{\rho_2} + 2m\sqrt{\lambda\sigma\epsilon}\sinh\omega, \qquad (50)\\
&\frac{\partial \log\sigma}{\partial u} = \left(2\sqrt{\frac{\lambda\epsilon}{\sigma}}\cosh\omega - \sqrt{E}\right)\frac{\alpha}{\lambda},\\
&\frac{\partial \log\sigma}{\partial v} = \left(2\sqrt{\frac{\lambda\epsilon}{\sigma}}\sinh\omega - \sqrt{G}\right)\frac{\beta}{\lambda},\\
&\frac{\partial \omega}{\partial u} = \frac{1}{\sqrt{E}}\frac{\partial \sqrt{G}}{\partial u} - \frac{\alpha}{\sqrt{\lambda\sigma\epsilon}}\sinh\omega,\\
&\frac{\partial \omega}{\partial v} = \frac{1}{\sqrt{G}}\frac{\partial \sqrt{E}}{\partial v} - \frac{\beta}{\sqrt{\lambda\sigma\epsilon}}\cosh\omega.
\end{aligned}
$$

In consequence of these equations, the expressions for $k$ and $l$ of the surface $\Sigma_1$, namely

$$
k = 2\sqrt{\frac{\lambda\epsilon}{\sigma}}\cosh\omega - \sqrt{E}, \qquad l = 2\sqrt{\frac{\lambda\epsilon}{\sigma}}\sinh\omega - \sqrt{G}, \qquad (51)
$$

satisfy the last two of equations (I). Moreover, it is readily seen that

$$
\frac{\partial}{\partial u}\left(\frac{1}{l}\frac{\partial k}{\partial v}\right) = \frac{\partial}{\partial v}\left(\frac{1}{k}\frac{\partial l}{\partial u}\right),
$$

so that $\Sigma_1$ is a surface satisfying (43), as was evident from general considerations.

From (26) and (39) we have

$$\sqrt{E'} = \sqrt{E} - \sqrt{\frac{\lambda\epsilon}{\sigma}}\,e^{-\omega}, \qquad -\sqrt{G'} = \sqrt{G} + \sqrt{\frac{\lambda\epsilon}{\sigma}}\,e^{-\omega}, \tag{52}$$

and from (25)

$$k' = \sqrt{\frac{\lambda\epsilon}{\sigma}}\,e^{\omega} - \sqrt{E}, \qquad -l' = \sqrt{\frac{\lambda\epsilon}{\sigma}}\,e^{\omega} - \sqrt{G}. \tag{53}$$

Equations (22) reduce to

$$\frac{1}{\sqrt{G'}}\frac{\partial\sqrt{E'}}{\partial v} + \frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v} - \frac{\beta}{\sqrt{\lambda\sigma\epsilon}}\,e^{-\omega} = 0,$$

$$\frac{1}{\sqrt{E'}}\frac{\partial\sqrt{G'}}{\partial u} + \frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u} + \frac{\alpha}{\sqrt{\lambda\sigma\epsilon}}\,e^{-\omega} = 0,$$

from which it readily follows that $E'$ and $G'$ satisfy (43),

Equations (21) become

$$\frac{\sqrt{E'}}{\rho_1'} = \frac{\sqrt{E}}{\rho_1} + \frac{\mu}{\sqrt{\lambda\sigma\epsilon}}\,e^{-\omega}, \qquad -\frac{\sqrt{G'}}{\rho_2'} = \frac{\sqrt{G}}{\rho_2} - \frac{\mu}{\sqrt{\lambda\sigma\epsilon}}\,e^{-\omega}. \tag{54}$$

From (2) it follows that if a transformation $R$ is determined by functions $\lambda$ and $\mu$, another $\bar{R}$, is given by $\lambda + c$ and $\mu$, where $c$ is any constant. Denoting the functions for $\bar{R}$ by $\bar{\lambda}, \bar{\mu}, \cdots, \bar{\sigma}$, we have from (I)

$$\bar{\lambda} = \lambda + c, \qquad \bar{\mu} = \mu, \qquad \bar{\alpha} = \alpha, \qquad \bar{\beta} = \beta, \qquad \bar{m} = m, \tag{55}$$

and from (6)

$$\bar{\lambda}\bar{\sigma} = \lambda\sigma, \qquad \bar{\sigma} = \frac{\lambda\sigma}{\lambda + c}.$$

Substituting these values in (I), we find that

$$\bar{k} = \frac{1}{\lambda}\left(k(\lambda + c) + c\sqrt{E}\right), \qquad \bar{l} = \frac{1}{\lambda}\left(l(\lambda + c) + c\sqrt{G}\right).$$

If $\bar{\Sigma}_1$ denotes the corresponding transform of $\Sigma$, it follows from (10) that $\Sigma_1$ and $\bar{\Sigma}_1$ correspond with parallelism of tangent planes.

For any transformation $R$ the tangent planes to $S$ are normal to the lines joining corresponding points on $\Sigma$ and $\Sigma_1$. From (5) and (55) it follows that $S$ and $\bar{S}$, the surface of centers of the transformation $\bar{R}$, correspond with parallelism of tangent planes.

When these results are applied to the type of transformations $R$ given by (50) we have

$$\bar{\omega} = \omega.$$

## 7. When $\Sigma$ is an isothermic surface

From (2) it follows that $\mu$ satisfies the tangential equation of $\Sigma$, that is the equation satisfied by the direction-cosines of the normal to $\Sigma$. Hence if $\overline{\Sigma}$ is any surface having the same spherical representation of its lines of curvature as $\Sigma$, the function $\mu$ and the function $\overline{\lambda}$, given by equations for $\overline{\Sigma}$ analogous to (2), determine a transformation $R$ for $\overline{\Sigma}$. We say that this transformation of $\overline{\Sigma}$ is obtained from the given transformation $R$ of $\Sigma$ by a *transformation of Combescure*. From this result it is clear that if we find all the transformations $R$ of the type (50) of isothermic surfaces, all the other solutions of our problem are obtainable from these by transformations of Combescure. Accordingly we confine ourselves hereafter to the case where $\Sigma$ is an isothermic surface.

When $\Sigma$ is isothermic, the parameters of the lines of curvature may be chosen so that

$$\sqrt{E} = \sqrt{G} = e^{\phi}. \tag{56}$$

Now

$$k + l = 2\left(\sqrt{\frac{\lambda\epsilon}{\sigma}}e^{\omega} - e^{\phi}\right). \tag{57}$$

From equations (50) we derive the equations

$$\frac{\partial}{\partial u}\left(\sqrt{\frac{\lambda\epsilon}{\sigma}}e^{\omega} - e^{\phi}\right) = \left(\sqrt{\frac{\lambda\epsilon}{\sigma}}e^{\omega} - e^{\phi}\right)\left(\frac{\partial\phi}{\partial u} - \frac{\alpha e^{\omega}}{\sqrt{\lambda\sigma\epsilon}}\right),$$

$$\frac{\partial}{\partial v}\left(\sqrt{\frac{\lambda\epsilon}{\sigma}}e^{\omega} - e^{\phi}\right) = \left(\sqrt{\frac{\lambda\epsilon}{\sigma}}e^{\omega} - e^{\phi}\right)\left(\frac{\partial\phi}{\partial v} - \frac{\beta e^{\omega}}{\sqrt{\lambda\sigma\epsilon}}\right),$$

and the last two of equations (50) reduce to

$$\frac{\partial\phi}{\partial u} - \frac{\alpha e^{\omega}}{\sqrt{\lambda\sigma\epsilon}} = \frac{\partial\omega}{\partial u} - \frac{\partial \log\sqrt{\lambda\sigma}}{\partial u},$$

$$\frac{\partial\phi}{\partial v} - \frac{\beta e^{\omega}}{\sqrt{\lambda\sigma\epsilon}} = \frac{\partial\omega}{\partial v} - \frac{\partial \log\sqrt{\lambda\sigma}}{\partial v}.$$

The integral of these equations is

$$e^{\omega} = \frac{\sqrt{\lambda\sigma\epsilon}e^{\phi}}{\lambda + c}, \tag{58}$$

where $c$ denotes a constant of integration.

8. TRANSFORMATIONS $D_m$ OF $\Sigma$. ISOTHERMIC SURFACES PARALLEL TO $\Sigma'$

When $c$ in (58) is zero, equations (50) become

$$
\begin{aligned}
&\frac{\partial\lambda}{\partial u} = e^{\phi}\,\alpha, \qquad \frac{\partial\lambda}{\partial v} = e^{\phi}\,\beta,\\
&\frac{\partial\mu}{\partial u} = -\frac{e^{\phi}}{\rho_1}\alpha, \qquad \frac{\partial\mu}{\partial v} = -\frac{e^{\phi}}{\rho_2}\beta,\\
&\frac{\partial\sigma}{\partial u} = e^{-\phi}\,\alpha, \qquad \frac{\partial\sigma}{\partial v} = -\,e^{-\phi}\,\beta,\\
&\frac{\partial\alpha}{\partial u} = -\frac{\partial\phi}{\partial v}\,\beta + \mu\frac{e^{\phi}}{\rho_1} + m(\sigma e^{\phi} + \lambda e^{-\phi}),\\
&\frac{\partial\alpha}{\partial v} = \beta\frac{\partial\phi}{\partial u}, \qquad \frac{\partial\beta}{\partial u} = \alpha\frac{\partial\phi}{\partial v},\\
&\frac{\partial\beta}{\partial v} = -\frac{\partial\phi}{\partial u}\,\alpha + \mu\frac{e^{\phi}}{\rho_2} + m(\sigma e^{\phi} - \lambda e^{-\phi}),
\end{aligned}
\tag{59}
$$

and from (51) we have

$$k = -\,l = \frac{\lambda}{\sigma}e^{-\phi}. \tag{60}$$

These are the equations of transformations $D_m$ of isothermic surfaces into isothermic surfaces discovered by Darboux.*

From (52) and (53) we obtain

$$\sqrt{E'} = e^{\phi} - \frac{\lambda}{\sigma}e^{-\phi}, \qquad -\sqrt{G'} = e^{\phi} + \frac{\lambda}{\sigma}e^{-\phi}, \tag{61}$$

and

$$k' = l' = 0. \tag{62}$$

Hence $\Sigma_1'$ is a point and not a surface.

Conversely, if a transformation $R$ can be deformed in such a way that one of the sheets of the new envelope is a point, the requirement that the lines of curvature on the two sheets of the new envelope shall correspond is satisfied identically and consequently the preceding formulas apply. For this case it follows from (53) that

$$\sqrt{E} = \sqrt{G} = \sqrt{\frac{\lambda\epsilon}{\sigma}}\,e^{\omega},$$

and consequently the transformation $R$ is a $D_m$.†

*Annales de l'École Normale Supérieure, ser. 3, vol. 16 (1899), pp. 491–508.

† Cf. Eisenhart, Rendiconti dei Lincei, l. c.

For the present case equations (20) become

$$\frac{\partial \log \rho}{\partial u} + \frac{\alpha}{\sigma} e^{-\phi} = 0, \qquad \frac{\partial \log \rho}{\partial v} - \frac{\beta}{\sigma} e^{-\phi} = 0,$$

from which it follows that $\rho\sigma$ is constant. Since the functions of a transformation $R$ are determined only to within a constant factor, it follows from § 2 that the functions by means of which $\Sigma'$ is transformed into the point $\Sigma_1'$ are given by

$$(63) \qquad \alpha' = \frac{\alpha}{\sigma}, \qquad \beta' = -\frac{\beta}{\sigma}, \qquad \mu' = \frac{\mu}{\sigma}, \qquad \lambda' = \frac{\lambda}{\sigma}, \qquad \sigma' = 1.$$

From the general results of this paper it follows that $\Sigma'$ has the same spherical representation of its lines of curvature as two isothermic surfaces, say $\overline{\Sigma}$ and $\overline{\Sigma}'$. From (54) it follows that this spherical representation is determined by

$$(64) \qquad \sqrt{\mathcal{E}'} = \frac{\sqrt{E'}}{\rho_1'} = \frac{e^{\phi}}{\rho_1} + \frac{\mu e^{-\phi}}{\sigma}, \qquad \sqrt{\mathcal{G}'} = \frac{\sqrt{G'}}{\rho_2'} = -\frac{e^{\phi}}{\rho_2} + \frac{\mu}{\sigma} e^{-\phi}.$$

Applying the results of § 5, we find that the first fundamental coefficients of $\overline{\Sigma}$ and $\overline{\Sigma}'$ are given by

$$(65) \qquad \sqrt{\overline{E}} = \sqrt{\overline{G}} = \sigma e^{-\phi}, \qquad \sqrt{\overline{E}'} = \sqrt{\overline{G}'} = e^{-\phi}/\sigma.$$

The principal radii of normal curvature of these surfaces are given by

$$(66) \qquad \begin{aligned} \frac{1}{\overline{\rho}_1} &= \frac{e^{-\phi}}{\sigma}\left(\frac{e^{\phi}}{\rho_1} + \frac{\mu}{\sigma} e^{-\phi}\right), & \frac{1}{\overline{\rho}_2} &= \frac{e^{-\phi}}{\sigma}\left(\frac{e^{\phi}}{\rho_2} - \frac{\mu}{\sigma} e^{-\phi}\right), \\ \frac{1}{\overline{\rho}_1'} &= e^{\phi}\sigma\left(\frac{e^{\phi}}{\rho_1} + \frac{\mu}{\sigma} e^{-\phi}\right), & \frac{1}{\overline{\rho}_2'} &= -e^{\phi}\sigma\left(\frac{e^{\phi}}{\rho_2} - \frac{\mu}{\sigma} e^{-\phi}\right).* \end{aligned}$$

## 9. Transformations $D_m$ of $\overline{\Sigma}$ and $\overline{\Sigma}'$, and their deformations

Since $\overline{\Sigma}$ and $\overline{\Sigma}'$ have the same spherical representation of their lines of curvature as $\Sigma'$, a transformation $R$ of each of these surfaces is given by taking for $\mu$ the value of $\mu'$ in (63). From the third and fourth of equations (I) we see that the corresponding functions $\alpha$ and $\beta$ differ at most in signs from $\alpha'$ and $\beta'$ as given by (63). We consider the two surfaces separately.

For $\overline{\Sigma}$ we find that the functions are given by

$$\bar{\alpha} = \frac{\alpha}{\sigma}, \qquad \bar{\beta} = \frac{\beta}{\sigma}, \qquad \mu = \frac{\mu}{\sigma}, \qquad \bar{\lambda} = \lambda + c,$$

$$\bar{\sigma} = \frac{-\lambda}{\sigma(\lambda + c)}, \qquad \bar{m} = -m,$$

* Evidently for the surfaces symmetric to $\overline{\Sigma}$ and $\overline{\Sigma}'$ with respect to a point, the expressions analogous to (65) are the same and to (66) differ in sign.

where $c$ is any constant. Furthermore,

$$\bar{k} = \left( ce^{\phi}\frac{\sigma}{\lambda} - (\lambda + c)e^{-\phi} \right), \qquad l = \left( ce^{\phi}\frac{\sigma}{\lambda} + (\lambda + c)e^{-\phi} \right).$$

When $c = 0$, the surface $\bar{\Sigma}_1$ is isothermic and the transformation is a $D_m$. The values of the functions for this case are

$$\alpha = \frac{\alpha}{\sigma}, \qquad \bar{\beta} = \frac{\beta}{\sigma}, \qquad \bar{\mu} = \frac{\mu}{\sigma}, \qquad \bar{\lambda} = \lambda, \qquad \bar{\sigma} = -\frac{1}{\sigma}. \tag{67}$$

From equations analogous to (12) we find for the principal radii of $\bar{\Sigma}_1$ the expressions

$$\frac{1}{\rho_{11}} = \frac{e^{2\phi}}{\lambda}\left(\frac{1}{\rho_1} + \frac{\mu}{\lambda}\right), \qquad \frac{1}{\rho_{12}} = \frac{e^{2\phi}}{\lambda}\left(\frac{1}{\rho_2} + \frac{\mu}{\lambda}\right). \tag{68}$$

In a similar manner, transformations of $\bar{\Sigma}'$ are given by

$$\alpha' = \frac{\alpha}{\sigma}, \qquad \bar{\beta}' = -\frac{\beta}{\sigma}, \qquad \bar{\mu}' = \frac{\mu}{\sigma}, \qquad \bar{\lambda}' = -\frac{1}{\sigma},$$

$$\bar{\sigma}' = \frac{\lambda}{1 - c\sigma}, \qquad \bar{m}' = -m,$$

$$\bar{k}' = \left( e^{\phi}\frac{c\sigma - 1}{\lambda} - ce^{-\phi} \right), \qquad \bar{l}' = -\left( e^{\phi}\frac{c\sigma - 1}{\lambda} + ce^{-\phi} \right).$$

In particular a transformation $D_{-m}$ of $\bar{\Sigma}'$ is given by

$$\alpha' = \frac{\alpha}{\sigma}, \qquad \bar{\beta}' = -\frac{\beta}{\sigma}, \qquad \bar{\mu}' = \frac{\mu}{\sigma}, \qquad \bar{\lambda}' = -\frac{1}{\sigma}, \qquad \bar{\sigma}' = \lambda, \tag{69}$$

and the fundamental functions for the transform $\bar{\Sigma}_1'$ have the expressions

$$\bar{k}' = -\bar{l}' = -e^{\phi}/\lambda,$$

$$\frac{1}{\bar{\rho}_{11}'} = \mu + \frac{\lambda}{\rho_1}, \qquad \frac{1}{\bar{\rho}_{12}'} = \mu + \frac{\lambda}{\rho_2}. \tag{70}$$

The isothermic surface defined by (70) has been considered by Bianchi;* he calls the process by which it is obtained from $\Sigma$ and the functions of a transformation $D_m$ of $\Sigma$ the transformation $T_m$ determined by the given $D_m$. This transformation was defined intrinsically by Bianchi without any indication of its geometrical relation to $\Sigma$.

We shall now apply the results of § 8 to these surfaces $\bar{\Sigma}$ and $\bar{\Sigma}'$, by deforming the surfaces of center $\bar{S}$ and $\bar{S}'$ of the transformations $D_{-m}$ of these surfaces defined by (67) and (69).

*Annali di matematica, ser. 3, vol. 12 (1906), pp. 19–54.

We call $(\Sigma')$ the surface analogous to $\Sigma'$. From (61) and (64) we find that its fundamental functions are given by

$$(\sqrt{E'}) = (\sigma e^{\phi} + \lambda e^{-\phi}), \qquad (\sqrt{G'}) = -(\sigma e^{\phi} - \lambda e^{-\phi}),$$

$$\left(\frac{\sqrt{E'}}{\rho_1'}\right) = \frac{e^{\phi}}{\rho_1}, \qquad \left(\frac{\sqrt{G'}}{\rho_2'}\right) = -\frac{e^{\phi}}{\rho_2}.$$

Hence $(\Sigma')$ has the same spherical representation of its lines of curvature as $\Sigma_1$. Furthermore, from equations analogous to (65) and (66), we find that the surfaces $(\bar{\Sigma})$, $(\bar{\Sigma}')$ analogous to $\bar{\Sigma}$, $\bar{\Sigma}'$ are defined by

$$(\sqrt{\bar{E}}) = (\sqrt{\bar{G}}) = e^{\phi}, \qquad \left(\frac{1}{\bar{\rho}_1}\right) = -\frac{1}{\rho_1}, \qquad \left(\frac{1}{\bar{\rho}_2}\right) = -\frac{1}{\rho_2},$$

$$(\sqrt{\bar{E}'}) = (\sqrt{\bar{G}'}) = e^{-\phi}, \qquad \left(\frac{1}{\bar{\rho}_1'}\right) = -\frac{e^{2\phi}}{\rho_1}, \qquad \left(\frac{1}{\bar{\rho}_2'}\right) = \frac{e^{2\phi}}{\rho_2}.$$

Hence $(\bar{\Sigma})$ and $(\Sigma')$ are respectively the symmetric of $\Sigma$ with respect to a point and its Christoffel transform.

Proceeding in like manner with $\bar{\Sigma}'$, we get for $(\Sigma')'$, the surface analogous to $\Sigma'$, the functions

$$(\sqrt{E'})' = \frac{1}{\lambda\sigma}(\sigma e^{\phi} + \lambda e^{-\phi}), \qquad (\sqrt{G'})' = \frac{1}{\lambda\sigma}(\sigma e^{\phi} - \lambda e^{-\phi}),$$

$$\left(\frac{\sqrt{E'}}{\rho_1'}\right)' = \left(\frac{e^{\phi}}{\rho_1} + \frac{\mu}{\sigma}e^{-\phi} + \frac{\mu}{\lambda}e^{\phi}\right), \qquad \left(\frac{\sqrt{G'}}{\rho_2'}\right)' = -\left(\frac{e^{\phi}}{\rho_2} + \frac{\mu}{\sigma}e^{-\phi} - \frac{\mu}{\lambda}e^{\phi}\right).$$

This surface has the same spherical representation of its lines of curvature as $\Sigma_1$. The surfaces $(\bar{\Sigma})'$ and $(\bar{\Sigma}')'$ are respectively the symmetric of $\Sigma_1$ and the Christoffel transform.

Let $S_0$ be the surface of centers of the spheres enveloped by $\Sigma'$ and $\Sigma_1'$, the latter being a point, say $O$. Corresponding normals to $\Sigma'$, $\bar{\Sigma}$, and $\bar{\Sigma}'$ are parallel. In like manner, the normals to $\bar{\Sigma}_1$ and $\bar{\Sigma}_1'$ are parallel to the lines joining $O$ to corresponding points of $S_0$. Hence the permanent conjugate system on $S_0$, that is the system corresponding to the lines of curvature on $\Sigma$, project upon the unit sphere with center $O$ into the orthogonal system representing the lines of curvature on $\bar{\Sigma}_1$ and $\bar{\Sigma}_1'$. Applying the same reasoning to the transformation resulting from the deformation of $\bar{S}$, the surface of centers of the spheres enveloped by $\bar{\Sigma}'$ and $\bar{\Sigma}_1'$, we see that the lines joining $O$ to points of $\bar{S}_0$, the deform of $\bar{S}$, are parallel to the normals to $\Sigma$. Furthermore, it is readily seen that $\bar{S}$ and $S_0$ correspond with parallelism of tangent planes; and likewise $S$ and $\bar{S}_0$.*

* The results of this paragraph were obtained by Bianchi by purely analytical processes as a result of his intrinsic definition of the transformation $T_m$. Cf. Rendiconti della R. Accademia dei Lincei, ser. 3, vol. 24 (1915), p. 386.

## 10. Case $c \neq 0$ in equation (58)

When we substitute the value (58) of $\omega$ with $c \neq 0$ in equations (50), they become

$$
\begin{aligned}
&\frac{\partial \lambda}{\partial u} = e^{\phi}\,\alpha, \qquad \frac{\partial \lambda}{\partial v} = e^{\phi}\,\beta,\\
&\frac{\partial \mu}{\partial u} = -\frac{e^{\phi}}{\rho_1}\alpha, \qquad \frac{\partial \mu}{\partial v} = -\frac{e^{\phi}}{\rho_2}\beta,\\
&\frac{\partial \sigma}{\partial u} = \frac{\alpha}{\lambda}\left(\frac{-ce^{\phi}\sigma}{\lambda + c} + (\lambda + c)e^{-\phi}\right),\\
&\frac{\partial \sigma}{\partial v} = -\frac{\beta}{\lambda}\left(\frac{ce^{\phi}\sigma}{\lambda + c} + (\lambda + c)e^{-\phi}\right),\\
&\frac{\partial \alpha}{\partial u} = -\frac{\partial \phi}{\partial v}\beta + \mu\frac{e^{\phi}}{\rho_1} + m\left(\frac{\lambda\sigma}{\lambda + c} + (\lambda + c)e^{-\phi}\right),\\
&\frac{\partial \alpha}{\partial v} = \beta\frac{\partial \phi}{\partial u}, \qquad \frac{\partial \beta}{\partial u} = \alpha\frac{\partial \phi}{\partial v},\\
&\frac{\partial \beta}{\partial v} = -\frac{\partial \phi}{\partial u}\alpha + \mu\frac{e^{\phi}}{\rho_2} + m\left(\frac{\lambda\sigma}{\lambda + c} - (\lambda + c)e^{-\phi}\right),
\end{aligned}
\tag{71}
$$

and

$$
k = -\frac{ce^{\phi}}{\lambda + c} + \frac{\lambda + c}{\sigma}e^{-\phi}, \qquad l = -\frac{ce^{\phi}}{\lambda + c} - \frac{\lambda + c}{\sigma}e^{-\phi}. \tag{72}
$$

For the surface $\Sigma'$ we have the functions

$$
\begin{aligned}
&\sqrt{E'} = e^{\phi} - \frac{\lambda + c}{\sigma}e^{-\phi}, \qquad -\sqrt{G'} = e^{\phi} + \frac{\lambda + c}{\sigma}e^{-\phi},\\
&\frac{\sqrt{E'}}{\rho_1'} = \frac{e^{\phi}}{\rho_1} + \frac{\mu}{\lambda}\frac{\lambda + c}{\sigma}e^{-\phi}, \qquad \frac{\sqrt{G'}}{\rho_2'} = -\frac{e^{\phi}}{\rho_2} + \frac{\mu}{\lambda}\frac{\lambda + c}{\sigma}e^{-\phi},
\end{aligned}
\tag{73}
$$

and for $\Sigma_1'$

$$
\begin{aligned}
&k' = -l' = -\frac{ce^{\phi}}{\lambda + c},\\
&\frac{1}{\rho_{11}'} = \frac{1}{c}\left(\frac{\lambda + c}{\rho_1} + \mu\right), \qquad \frac{1}{\rho_{12}'} = \frac{1}{c}\left(\frac{\lambda + c}{\rho_2} + \mu\right).
\end{aligned}
\tag{74}
$$

The last two are a consequence of (73) and of equations analogous to (12). From these expressions it is seen that $\Sigma_1'$ is an isothermic surface. We shall now determine its relation to $\Sigma$.

It is readily seen that, if $\lambda$, $\mu$, $\alpha$, $\beta$, $\sigma$ are solutions of equations (71), a

set of solutions of (59) is given by

$$(\lambda) = \lambda + c, \qquad (\mu) = \mu, \qquad (\alpha) = \alpha, \qquad (\beta) = \beta,$$

$$(\sigma) = \frac{\lambda\sigma}{\lambda + c}, \qquad (m) = m.$$

Hence these functions define a transformation $(D_m)$ of $\Sigma$, and a comparison of (70) and (74) reveals the fact that $\Sigma_1'$, as given by (74) is homothetic to the transform of $\Sigma$ by the $T_m$ determined by $(D_m)$.

## 11. When $\Sigma_1'$ is a plane

From equations (21), (53), and equations for $\Sigma'$ analogous to (12), we find the following expressions for the principal radii of normal curvature of $\Sigma_1'$:

$$\begin{aligned} &\frac{1}{\rho_{11}'}\left(\sqrt{\frac{\lambda\epsilon}{\sigma}}e^{\omega} - \sqrt{E}\right) + \frac{\sqrt{E}}{\rho_1} + \frac{\mu}{\sqrt{\lambda\sigma\epsilon}}e^{\omega} = 0, \\ &\frac{1}{\rho_{12}'}\left(\sqrt{\frac{\lambda\epsilon}{\sigma}}e^{\omega} - \sqrt{G}\right) + \frac{\sqrt{G}}{\rho_2} + \frac{\mu}{\sqrt{\lambda\sigma\epsilon}}e^{\omega} = 0. \end{aligned} \tag{75}$$

Hence the necessary condition that $\Sigma_1'$ be a plane is

$$\frac{\sqrt{E}}{\rho_1} = \frac{\sqrt{G}}{\rho_2} = -\frac{\mu}{\sqrt{\lambda\sigma\epsilon}}e^{\omega} = e^{-\theta}, \tag{76}$$

where $\theta$ is a function thus defined. This condition is also sufficient; otherwise $k'$, $l'$, or both, is zero and the lines of curvature on $\Sigma_1'$ are minimal.

In consequence of (76) equations (50) become

$$\begin{aligned} &\frac{\partial\lambda}{\partial u} = \sqrt{E}\alpha, \qquad \frac{\partial\lambda}{\partial v} = \sqrt{G}\beta, \\ &\frac{\partial\mu}{\partial u} = -e^{-\theta}\alpha, \qquad \frac{\partial\mu}{\partial v} = -e^{-\theta}\beta, \\ &\frac{\partial\alpha}{\partial u} = \frac{\partial\theta}{\partial v}\beta + \mu e^{-\theta} - m\left(\frac{\lambda\sigma}{\mu}e^{-\theta} + \mu e^{\theta}\right), \\ &\frac{\partial\alpha}{\partial v} = -\beta\frac{\partial\theta}{\partial u}, \qquad \frac{\partial\beta}{\partial u} = -\alpha\frac{\partial\theta}{\partial v}, \\ &\frac{\partial\beta}{\partial v} = \frac{\partial\theta}{\partial u}\alpha + \mu e^{-\theta} - m\left(\frac{\lambda\sigma}{\mu}e^{-\theta} - \mu e^{\theta}\right), \\ &\frac{\partial}{\partial u}\frac{\sigma\lambda}{\mu} = -\alpha e^{\theta}, \qquad \frac{\partial}{\partial v}\frac{\sigma\lambda}{\mu} = \beta e^{\theta}, \end{aligned} \tag{77}$$

and the last two of equations (50) are satisfied identically. Now

$$(78)\quad k = -\left(e^{-\theta}\frac{\lambda}{\mu} + e^{\theta}\frac{\mu}{\sigma} + \sqrt{E}\right), \qquad l = -\left(e^{-\theta}\frac{\lambda}{\mu} - e^{\theta}\frac{\mu}{\sigma} + \sqrt{G}\right).$$

Making use of (11) and (12), we get

$$(79)\quad \frac{\sqrt{E_1}}{\rho_{11}} = \frac{\sqrt{G_1}}{\rho_{12}} = -e^{\theta}\frac{\mu^2}{\lambda\sigma}.$$

Hence $\Sigma$ and $\Sigma_1$ are surfaces with isothermal spherical representation of their lines of curvature. Equations (77) define the transformations $E_m$ of surfaces with isothermal representation of their lines of curvature into surfaces of the same kind, which we have discussed previously from several points of view.*

From (53) and (54) we get

$$\sqrt{E'} = \frac{\mu e^{\theta}}{\sigma} + \sqrt{E}, \qquad \sqrt{G'} = \frac{\mu e^{\theta}}{\sigma} - \sqrt{G},$$

$$(80)\quad \sqrt{\mathcal{E}'} = \frac{\sqrt{E'}}{\rho_1'} = e^{-\theta} - \frac{\mu^2}{\lambda\sigma}e^{\theta},$$

$$\sqrt{\mathcal{G}'} = \frac{\sqrt{G'}}{\rho_2'} = -e^{-\theta} - \frac{\mu^2}{\lambda\sigma}e^{\theta},$$

$\mathcal{E}'$ and $\mathcal{G}'$ being the coefficients of the spherical representation of $\Sigma'$. Also

$$(81)\quad k' = -\left(\frac{\lambda}{\mu}e^{-\theta} + \sqrt{E}\right), \qquad l' = \frac{\lambda}{\mu}e^{-\theta} - \sqrt{G},$$

so that the parametric curves on the plane $\Sigma_1'$ form an orthogonal system. Hence *the permanent conjugate system on* $S_0$, *the locus of centers of the spheres tangent to* $\Sigma'$ *and* $\Sigma_1'$, *projects into an orthogonal system on the plane* $\Sigma_1'$.†

## 12. Isothermic surfaces determined by a transformation $E_m$

From the general theory of § 5 we know that there are two isothermic surfaces $\overline{\Sigma}$ and $\overline{\Sigma}'$ with the same spherical representation of their lines of curvature as $\Sigma'$. We proceed to their determination.

From (80) and (49) we find that the first fundamental coefficients of these surfaces are

$$(82)\quad \sqrt{\overline{E}} = \sqrt{\overline{G}} = \frac{\mu}{\lambda\sigma}e^{\theta}, \qquad \sqrt{\overline{E}'} = \sqrt{\overline{G}'} = \frac{\lambda\sigma}{\mu}e^{-\theta},$$

*Annals of Mathematics, ser. 2, vol. 17 (1915), p. 64.

† Cf. Bianchi, Rendiconti dei Lincei, ser. 5, vol. 24 (1915), p. 378.

and their principal radii have the values

$$\frac{1}{\bar{\rho}_1} = \frac{\lambda\sigma}{\mu} e^{-2\theta} - \mu, \qquad \frac{1}{\bar{\rho}_2} = -\frac{\lambda\sigma}{\mu} e^{-2\theta} - \mu,$$

$$\frac{1}{\bar{\rho}_1'} = \frac{\mu}{\lambda\sigma} - \frac{\mu^3}{\lambda^2\sigma^2} e^{2\theta}, \qquad \frac{1}{\bar{\rho}_2'} = \frac{\mu}{\lambda\sigma} + \frac{\mu^3}{\lambda^2\sigma^2} e^{2\theta}. \tag{83}$$

From (20), (80), and (78) it follows that we may take $\rho = \mu/\lambda\sigma$, so that the transformation $R$ of $\Sigma'$ into $\Sigma_1'$ is given by

$$\alpha' = \frac{\alpha\mu}{\lambda\sigma}, \qquad \beta' = -\frac{\beta\mu}{\lambda\sigma}, \qquad \mu' = \frac{\mu^2}{\lambda\sigma}, \qquad \lambda' = \frac{\mu}{\sigma}, \qquad \sigma' = \frac{\mu}{\lambda}.$$

The transformations $R$ of $\bar{\Sigma}$ which are the Combescure transforms of the former are defined by

$$\bar{\alpha} = \frac{\alpha\mu}{\lambda\sigma}, \qquad \bar{\beta} = -\frac{\beta\mu}{\lambda\sigma}, \qquad \bar{\mu} = \frac{\mu^2}{\lambda\sigma},$$

$$\bar{\lambda} = \frac{\mu}{\lambda\sigma} + c, \qquad \bar{\sigma} = \frac{-\mu}{1 + c\dfrac{\lambda\sigma}{\mu}}, \qquad \bar{m} = -m, \tag{84}$$

where $c$ denotes a constant. From geometrical considerations we know that $\bar{\Sigma}_1$ is a plane; it is referred to an orthogonal system whose coefficients are

$$\bar{k} = ce^{\theta} - \frac{e^{-\theta}}{\mu}\left(1 + \frac{c\lambda\sigma}{\mu}\right), \qquad \bar{l} = ce^{\theta} + \frac{e^{-\theta}}{\mu}\left(1 + \frac{c\lambda\sigma}{\mu}\right). \tag{85}$$

When $c = 0$, the coördinate system on the plane $\bar{\Sigma}_1$ is isothermal. Moreover, the functions (84) with $c = 0$ satisfy equations (59) and determine a transformation $D_{-m}$ of $\bar{\Sigma}$.

In like manner we have transformations $R$ of $\bar{\Sigma}'$ given by

$$\bar{\alpha}' = \frac{\alpha\mu}{\lambda\sigma}, \qquad \bar{\beta}' = \frac{\beta\mu}{\lambda\sigma}, \qquad \bar{\mu}' = \frac{\mu^2}{\lambda\sigma},$$

$$\bar{\lambda}' = c - \mu, \qquad \bar{\sigma}' = \frac{-\mu^2}{\lambda\sigma(c - \mu)}, \qquad \bar{m}' = -m,$$

and

$$\bar{k}' = -ce^{-\theta}\frac{\lambda\sigma}{\mu^2} + e^{\theta}(c - \mu),$$

$$\bar{l}' = -ce^{-\theta}\frac{\lambda\sigma}{\mu^2} - e^{\theta}(c - \mu).$$

The surface $\bar{\Sigma}_1'$ is a plane, and for $c = 0$ we have a transformation $D_{-m}$ of $\bar{\Sigma}'$.

The isothermic surfaces $\bar{\Sigma}$ and $\bar{\Sigma}'$ belong to the group discussed in § 8. For every transformation $E_m$ has for one of its Combescure transforms a $D_m$, namely the transformation of the minimal surface having the given isothermal spherical representation of its lines of curvature.

PRINCETON UNIVERSITY

# JACOBI'S CONDITION FOR THE PROBLEM OF LAGRANGE IN THE CALCULUS OF VARIATIONS*

BY

DAVID M. SMITH

## INTRODUCTION

The problem of Lagrange, in the calculus of variations, is that of minimizing an integral

$$J = \int_{x_1}^{x_2} f(x, y_1, \cdots, y_n, y_1', \cdots, y_n')\, dx \tag{1}$$

with respect to curves which join two fixed points, 1 and 2, and satisfy a system of differential equations of the form

$$\phi_\beta(x, y_1, \cdots, y_n, y_1', \cdots, y_n') = 0 \qquad (\beta = 1, 2, \cdots, m, m < n). \tag{2}$$

For this problem proofs of the necessity of the Euler-Lagrange Rule and the conditions of Weierstrass and Legendre are obtainable without the use of the second variation.† For the Jacobi condition, however, the situation is less satisfactory. Kneser's proof‡ of this condition, which is based on the theory of envelopes, excludes important special cases, while the proofs that utilize the second variation, though more inclusive,§ involve elaborate and complicated transformations. Similar remarks apply to the derivations of the Legendre condition by Clebsch|| and von Escherich,¶ both proofs being based on transformations of the second variation. Hahn's deduction** of the Legendre condition from that of Weierstrass is more direct.

The object of this paper is to secure inclusive proofs of the necessity of the Legendre and Jacobi conditions by means of the second variation, but without the use of complicated transformations. If the curves of the family

* Presented to the Society, April 21, 1916.

† Bolza, *Vorlesungen über Variationsrechnung*, pp. 558, 603.

‡ Bolza, loc. cit., p. 610.

§ In this connection see remarks by Bolza, loc. cit., p. 634.

|| Journal für Mathematik, vol. 55 (1858), p. 254.

¶ Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften zu Wien, vol. 107 (1898), p. 1267.

** Monatshefte für Mathematik und Physik, vol. 17 (1906), p. 295. See also Bolza, loc. cit., p. 607.

(3) $$y_i = y_i(x, e) \qquad (i = 1, 2, \cdots, n),$$

pass through the points 1 and 2, satisfy equations (2), and contain, for the parameter value $e = 0$, the minimizing curve $E_{12}$,

$$y_i = y_i(x, 0),$$

then the function

(4) $$J(e) = \int_{x_1}^{x_2} f(x, y(x, e), y'(x, e))\, dx$$

must have a minimum for $e = 0$. The arc $E_{12}$ must be an extremal, and there exists a set of functions $\lambda_\beta(x)$ $(\beta = 1, 2, \cdots, m)$, such that if

$$F = f + \sum_\beta \lambda_\beta \phi_\beta,$$

then the derivatives $J'(0)$ and $J''(0)$ are expressible in the forms*

$$J'(0) = \int_{x_1}^{x_2} \sum_i (\eta_i F_{y_i} + \eta_i' F_{y_i'})\, dx,$$

$$J''(0) = \int_{x_1}^{x_2} \sum_{ik} (\eta_i \eta_k F_{y_i y_k} + 2\eta_i \eta_k' F_{y_i y_k'} + \eta_i' \eta_k' F_{y_i' y_k'})\, dx,$$

where the functions $\eta_i(x)$ are the variations of the family (3), defined by the relations

(5) $$\left(\frac{\partial y(x, e)}{\partial e}\right)_{e=0} = \eta_i(x).$$

If the arc $E_{12}$ minimizes the integral (1) the first variation $J'(0)$ must vanish, and the condition

(6) $$J''(0) \geqq 0$$

must also be satisfied, at least in the normal case, for every set of functions $\eta_i(x)$ which satisfy the equations

(7) $$\Phi_\beta(x, \eta_1, \cdots, \eta_n, \eta_1', \cdots, \eta_n') = \sum_i \left(\eta_i \frac{\partial \phi_\beta}{\partial y_i} + \eta_i' \frac{\partial \phi_\beta}{\partial y_i'}\right) = 0,$$

(8) $$\eta_i(x_1) = \eta_i(x_2) = 0.$$

The condition (6) implies a problem of Lagrange in the $x\eta$-space of precisely the same type as the original problem in the $xy$-space. The integral to be minimized is $J''(0)$, and the equations of condition corresponding to equations (2) are the equations (7). To this $x\eta$-problem we apply the Euler-Lagrange Rule, the Weierstrass condition, and the corner-point condition,† all of which would necessarily be included in a complete discussion of the

* Bolza, loc. cit., p. 558 and p. 620.

† Bolza, loc. cit., p. 571.

Lagrange problem, and which, as was remarked above, may be derived without the use of the second variation. The application of the Weierstrass condition to the $x\eta$-problem just described furnishes a proof of the Legendre condition for the $xy$-problem which is even simpler than that of Hahn. The Euler-Lagrange Rule and the corner-point condition lead to a proof of the Jacobi condition which includes the exceptional cases of Kneser's method, and which is much more simple and direct than the proofs which are based on the usual elaborate transformations of the second variation.

## 1. Preliminary definitions and theorems

In this paper the following notations will be used:

$$\frac{\partial F}{\partial y_i} = F_i, \qquad \frac{\partial F}{\partial y_i'} = G_i, \qquad \frac{\partial \phi_\beta}{\partial y_i} = \phi_{\beta i}, \qquad \frac{\partial \phi_\beta}{\partial y_i'} = \psi_{\beta i},$$

$$\frac{\partial^2 F}{\partial y_i \partial y_k} = P_{ik}, \qquad \frac{\partial^2 F}{\partial y_i \partial y_k'} = Q_{ik}, \qquad \frac{\partial^2 F}{\partial y_i' \partial y_k'} = R_{ik}.$$

The various indices $\alpha, i, r$, etc., will always have the ranges:

$$\alpha, \beta = 1, 2, \cdots, m; \qquad i, j, k = 1, 2, \cdots, n; \qquad r, s = 1, 2, \cdots, n - m;$$

$$h = 1, 2, \cdots, 2n; \qquad \nu = 1, 2, \cdots, 2n + m;$$

and the integral (1) and the equations (2) and (7) will be written in the abbreviated forms

$$J = \int_{x_1}^{x_2} f(x, y, y')\, dx; \qquad \phi_\beta(x, y, y') = 0; \qquad \Phi_\beta(x, \eta, \eta') = 0.$$

Concerning the functions $f$, $\phi_\beta$, and the arc $E_{12}$ the following hypotheses are made:

(*A*) *The arc $E_{12}$ is of class* $C''$ throughout the interval $(x_1, x_2)$, and passes through two fixed points* 1 *and* 2.

(*B*) *The functions $f(x, y, y')$, $\phi_\beta(x, y, y')$ are continuous and possess continuous partial derivatives of the first three orders in a region $T$ of points $(x, y, y')$, containing the values $(x, y, y')$ along $E_{12}$ in its interior.*

(*C*) *The arc $E_{12}$ minimizes the integral* (1) *with respect to curves which lie entirely within the region $T$, pass through the points* 1 *and* 2, *and satisfy the differential equations* (2).

(*D*) *The arc $E_{12}$ is normal in every sub-interval $(\xi_1, \xi_2)$ of $(x_1, x_2)$, i. e., the only system of $m$ functions $\lambda_\beta(x)$ of class $C'$ in $(\xi_1, \xi_2)$, which satisfies the equations*

* A curve is said to be of class $C^{(n)}$ if its coördinates, $y_i(x)$, are continuous and possess continuous derivatives of the first $n$ orders.

$$\sum_{\beta}\left(\lambda_\beta\,\phi_{\beta i}-\frac{d}{dx}(\lambda_\beta\,\psi_{\beta i})\right)=0,$$

*is* $\lambda_\beta = 0$.

(*E*) *At every point of the arc* $E_{12}$ *the condition*

$$|\psi_{\beta\alpha}| \neq 0$$

*is satisfied.*

An *admissible curve* is defined as a curve of class $D'$,* which lies entirely within the region $T$, passes through the points 1 and 2, and satisfies equations (2). The *variations* of a one parameter family (3) are the functions $\eta_i(x)$ defined by equations (5). A set of functions $\eta_i(x)$ which satisfy equations (7) and (8) is called a set of *admissible variations*, and it may be proved† as a result of the hypotheses just made that the elements $\eta_i(x)$ of every solution of equations (7) and (8) are the variations of a one-parameter family (3).

Let the family (3) be a set of admissible curves containing the minimizing arc $E_{12}$ for the parameter value $e = 0$. Substituting the functions $y_i(x, e)$ in the integral (1) we obtain the function $J(e)$ of equation (4). In terms of the notations described above the conditions on the derivatives $J'(O)$ and $J''(O)$ are expressible in the forms

$$(9)\qquad J'(O)=\int_{x_1}^{x_2}\sum_i(\eta_i F_i+\eta_i' G_i)\,dx=0,$$

$$J''(O)=\int_{x_1}^{x_2}\sum_{ik}(\eta_i\,\eta_k P_{ik}+2\eta_i\,\eta_k' Q_{ik}+\eta_i'\,\eta_k' R_{ik})\,dx\geqq 0,$$

and these equations must be satisfied for every set of admissible variations $\eta_i(x)$. A consequence of the vanishing of the first variation (9) is the well-known Euler-Lagrange Rule:

*If* $E_{12}$ *is a minimizing arc there exists a set of multipliers* $\lambda_\beta(x)$ *such that at every point of* $E_{12}$ *the equations*

$$(10)\qquad F_i-\frac{d}{dx}G_i=0$$

*are satisfied.*

An arc of class $C''$, whose coördinates $y_i$ with $m$ functions $\lambda_\beta(x)$, of class $C'$, satisfy equations (2) and (10), is called an *extremal*. Under the hypotheses made the following theorems may be proved:‡

THEOREM 1. *If* $x'$ *and* $x''$ *are the abscissas of two arbitrary points* $P'$ *and* $P''$ *on the arc* $E_{12}$, *then an n parameter family of curves*

* A curve of class $D^{(n)}$ is continuous and is composed of a finite number of arcs, each of which is of class $C^{(n)}$.

† Bolza, loc. cit., p. 568.

‡ Bolza, loc. cit., pp. 604, 565.

$$y_i = y_i(x, b_1, \cdots, b_n)$$

*can always be found, with the following properties:*

(*a*) *The family contains the arc* $E_{12}$ *for the values* $(b_1, \cdots, b_n) = (0, \cdots, 0)$, $x_1 \leqq x \leqq x_2$, *as indicated by the conditions*

$$y_i(x, 0, \cdots, 0) = y_i(x).$$

(*b*) *The functions* $y_i$, $y_i'$, $\partial y_i/\partial b_k$, $\partial^2 y_i/\partial x \partial b_k$ *are continuous within a certain neighborhood of the values* $(x, b_1, \cdots, b_n)$ *defining the arc* $E_{12}$.

(*c*) *The curves all pass through the point* $P'$, i. e.,

$$y_i(x', b_1, \cdots, b_n) = y_i(x').$$

(*d*) *The functions* $y_i(x, b_1, \cdots, b_n)$ *satisfy equations* (2).

(*e*) *For the values* $x = x''$, $(b_1, \cdots, b_n) = (0, \cdots, 0)$, *defining the point* $P''$ *the condition*

$$\left|\frac{\partial y_i}{\partial b_k}\right| \neq 0$$

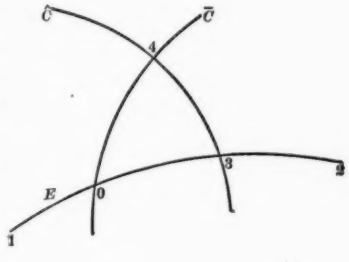
*is satisfied.*

THEOREM 2. *To every normal extremal arc of class* $C''$ *there belongs an unique set of multipliers* $\lambda_\beta(x)$.

With the aid of Theorem 1 the Weierstrass condition* may now be derived. Through an arbitrary point 3 on the arc $E_{12}$ a curve $\hat{C}$ is drawn,

$$y_i = \hat{y}_i(x),$$

whose coördinates are of class $C'$ and satisfy equations (2). The values $\hat{y}_i'(x_3)$ may be assigned arbitrarily, except that the set of values $x_3$, $\hat{y}_i(x_3)$, $\hat{y}_i'(x_3)$ must lie entirely within the region $T$ and satisfy condition $(E)$. Let 4 be a point on the curve $\hat{C}$ whose abscissa is $x_4 = x_3 - e$, $(e > 0)$. By Theorem 1 we can draw through a point $O$ on the arc $E_{12}$ a curve $\bar{C}$ of class $C'$,




$$y_i = \bar{y}_i(x, e),$$

which satisfies equations (2), passes through the point 4,

$$\bar{y}_i(x_4, e) = \hat{y} \quad x_4),$$

and reduces to the arc $E_{12}$ for the value $e = 0$.

The one-parameter family $V$,

$$V = E_{10} + \bar{C}_{04} + \hat{C}_{43} + E_{32},$$

* Bolza, loc. cit., p. 603.

is a set of admissible curves. Since all the arcs satisfy equations (2), $f$ may be replaced by $F$, and the function $J(e)$ for the family $V$ expressed in the form

$$J(e) = \int_{x_1}^{x_0} F dx + \int_{x_0}^{x_4} \bar{F} dx + \int_{x_4}^{x_3} \hat{F} dx + \int_{x_3}^{x_2} F dx,$$

where the multipliers $\lambda_\beta(x)$ are those belonging to the arc $E_{12}$. Now $J(e)$ must be an increasing function for the value $e = 0$, and consequently the condition $J'(O) \geqq 0$ must be satisfied. On evaluating $J'(O)$ with the aid of partial integration and equations (10) the following condition, due to Weierstrass, is obtained:

*If the arc $E_{12}$ minimizes the integral* (1) *the condition*

$$E(x, y(x), y'(x), p, \lambda) = F(x, y(x), p, \lambda) - F(x, y(x), y'(x), \lambda) - \sum_i (p_i - y_i') G_i(x, y(x), y'(x), \lambda) \geqq 0$$

*must be satisfied at every point $P$ of $E_{12}$ for all values $p_i = \hat{y}_i'$ belonging to an admissible arc through $P$.*

## 2. A NECESSARY CONDITION AT A CORNER POINT

A continuous curve $y_i = y_i(x)$ is said to have a *corner point* at $x = x_3$ if the condition

$$y_i'(x_3 - 0) \neq y_i'(x_3 + 0)$$

is satisfied for at least one value of $i$. To derive a necessary condition that a curve with a corner point shall minimize the integral (1) suppose given two normal extremal arcs of class $C''$:

$$E_{13} : y_i = y_i(x), \qquad E_{32} : y_i = \bar{y}_i(x),$$

which join the fixed points 1 and 3, and 3 and 2 respectively. Let the curve $E_{12}$ composed of the arcs $E_{13}$ and $E_{32}$ minimize the integral (1) with respect to admissible curves.

By Theorem 1 an $n$-parameter family of curves can be found,

$$y_i = y_i(x, b_1, \cdots, b_n), \tag{11}$$

which are of class $C'$, pass through the point 1, contain the arc $E_{12}$ for the values $(b_1, \cdots, b_n) = (0, \cdots, 0)$, satisfy equations (2), and for which the condition

$$\left| \frac{\partial y_i}{\partial b_k} \right| \neq 0 \tag{12}$$

is satisfied for the values $x = x_3$, $(b_1, \cdots, b_n) = (0, \cdots, 0)$ defining the point 3.

The parameters $b_k$ may be determined as functions of new parameters $e_i$ so that the conditions

$$y_i(x_3, b_1, \cdots, b_n) = e_i + y_i(x_3) \tag{13}$$

are satisfied. For, corresponding to the values $(e_1, \cdots, e_n) = (0, \cdots, 0)$ there is a solution, $b_i = 0$, of equations (13). Since the condition (12) is satisfied for the values $x = x_3$, $(b_1, \cdots, b_n) = (0, \cdots, 0)$ it is possible to solve equations (13) for the $b_i$ and obtain $n$ functions

$$b_i = b_i(e_1, \cdots, e_n), \tag{14}$$

which are of class $C'$ within a certain neighborhood $|e_k| < d$ and satisfy the conditions $b_i(0, \cdots, 0) = 0$. Substitution of the functions (14) in equations (11) gives an $n$-parameter family of curves

$$y_i = y_i(x, e_1, \cdots, e_n), \tag{15}$$

which pass through the point 1,

$$y_i(x_1, e_1, \cdots, e_n) = y_i(x_1), \tag{16}$$

contain the arc $E_{12}$ for the values $(e_1, \cdots, e_n) = (0, \cdots, 0)$, and satisfy the conditions

$$y_i(x_3, e_1, \cdots, e_n) = e_i + y_i(x_3). \tag{17}$$

In the same manner we construct a second family of curves

$$y_i = \bar{y}_i(x, e_1, \cdots, e_n), \tag{18}$$

which pass through the point 2,

$$\bar{y}_i(x_2, e_1, \cdots, e_n) = \bar{y}_i(x_2), \tag{19}$$

contain the arc $E_{12}$ for the values $(e_1, \cdots, e_n) = (0, \cdots, 0)$, and satisfy the conditions

$$\bar{y}_i(x_3, e_1, \cdots, e_n) = e_i + y_i(x_3). \tag{20}$$

The arcs (15) and (18) form an $n$ parameter family of admissible curves. It is therefore permissible to replace $f$ by $F$ and consider the integral

$$J(e_1, \cdots, e_n) = \int_{x_1}^{x_3} F(x, y(x, e), y'(x, e), \lambda)\, dx + \int_{x_3}^{x_2} F(x, \bar{y}(x, e), \bar{y}'(x, e), \bar{\lambda})\, dx,$$

in which the functions $\lambda_\beta(x)$ and $\bar{\lambda}_\beta(x)$ are the multipliers belonging to $E_{13}$ and $E_{32}$ respectively. Since $J(0, \cdots, 0)$ must be a minimum the conditions

$$(21)\quad \frac{\partial J}{\partial e_k} = \int_{x_1}^{x_2} \sum_i (\eta_{ik} F_i + \eta'_{ik} G_i)\, dx + \int_{x_2}^{x_3} \sum_i (\bar{\eta}_{ik} F_i + \bar{\eta}'_{ik} G_i)\, dx = 0$$

must be satisfied, where

$$\eta_{ik} = \frac{\partial y_i}{\partial e_k}, \qquad \bar{\eta}_{ik} = \frac{\partial \bar{y}_i}{\partial e_k}, \qquad (e_1, \cdots, e_n) = (0, \cdots, 0).$$

Apply now the usual integration by parts and use equations (10). Equations (21) then become

$$(22)\quad \sum_i \eta_{ik} G_i \Big]_{x_1}^{x_2} + \sum_i \bar{\eta}_{ik} G_i \Big]_{x_2}^{x_3} = 0.$$

But from equations (16) and (19)

$$\eta_{ik}(x_1) = \bar{\eta}_{ik}(x_2) = 0,$$

and from equations (17) and (20)

$$\eta_{ik}(x_3) = \bar{\eta}_{ik}(x_3) = \delta_{ik},$$

where $\delta_{ik}$ is zero for $i \neq k$ and unity for $i = k$. On substituting these values in equations (22) we obtain

$$G_i(x, y, y', \lambda)|_{x_3} = G_i(x, \bar{y}, \bar{y}', \bar{\lambda})|_{x_3}.$$

The following necessary condition at a corner point may therefore be stated:

*If the minimizing curve $E_{12}$ has a corner point at $x = x_3$ the conditions*

$$G_i(x_3 - 0) = G_i(x_3 + 0) \qquad (i = 1, 2, \cdots, n)$$

*must be satisfied.*

### 3. The Legendre condition

If we write

$$(23)\quad \omega(x, \eta, \eta') = \sum_{ik} (\eta_i \eta_k P_{ik} + 2\eta_i \eta'_k Q_{ik} + \eta'_i \eta'_k R_{ik}),$$

then the condition

$$(24)\quad J''(O) = \int_{x_1}^{x_2} \omega(x, \eta, \eta')\, dx \geqq 0$$

must be satisfied for every set of admissible variations $\eta_i(x)$. The $x\eta$-problem implied in this condition is of precisely the same type as the original problem in $xy$-space. In the first place it is clear from equations (23) and (7) that $\omega$ and $\Phi_\beta$, as functions of $x, \eta, \eta'$, satisfy the continuity conditions $(B)$. Furthermore, since

$$\sum_i \left[ \lambda_\beta \frac{\partial \Phi_\beta}{\partial \eta_i} - \frac{d}{dx}\left( \lambda_\beta \frac{\partial \Phi_\beta}{\partial \eta'_i} \right) \right] = \sum_i \left[ \lambda_\beta \phi_{\beta i} - \frac{d}{dx} (\lambda_\beta \psi_{\beta i}) \right],$$

it follows from hypothesis $(D)$ that every extremal of the $x\eta$-problem is normal throughout the interval $(x_1, x_2)$. We may therefore apply to the $x\eta$-problem the theory previously developed.

By the Euler-Lagrange rule the coördinates $\eta_i$ of a curve which minimizes the integral $J''(O)$ must satisfy, with $m$ functions $\mu_\beta$ of class $C'$, the equations

$$\Omega_{\eta_i} - \frac{d}{dx}\Omega_{\eta_i'} = 0, \tag{25}$$

where

$$\Omega = \omega + \sum_\beta \mu_\beta \Phi_\beta .$$

Equations (25), when written out, form with equations (7) the system

$$\begin{aligned} &\sum_k \left[ \eta_k P_{ik} + \eta_k' Q_{ik} - \frac{d}{dx}(\eta_k Q_{ki} + \eta_k' R_{ik}) \right] \\ &\qquad + \sum_\beta \left[ \mu_\beta \phi_{\beta i} - \frac{d}{dx}(\mu_\beta \psi_{\beta i}) \right] = 0, \\ &\sum_i [\phi_{\beta i}\eta_i + \psi_{\beta i}\eta_i'] = 0. \end{aligned} \tag{26}$$

Equations $(26_1)$, which are linear and homogeneous in $\eta_i$, $\mu_\beta$ and their derivatives, are known in the theory of the second variation as the Jacobi equations. It will be understood hereafter that the elements $u_i(x)$, $\rho_\beta(x)$ of a solution $(u; \rho)$ of equations (26) are of class $C''$ and $C'$ respectively.

If now the condition (24) is satisfied the arc in $x\eta$-space,

$$C: \qquad \eta_i \equiv 0, \qquad x_1 \leqq x \leqq x_2,$$

must be a minimizing extremal for the integral $J''(O)$, since its coördinates $\eta_i$, with the multipliers $\mu_\beta \equiv 0$, satisfy equations (26) and give $J''(O)$ its minimum value zero. The uniqueness of the set of multipliers $\mu_\beta \equiv 0$ is a consequence of hypothesis $(D)$. Since the arc $C$ minimizes the integral $J''(O)$ in the $x\eta$-space the Weierstrass condition for the $x\eta$-problem,

$$E(x, \eta, \eta', \zeta, \mu) \geqq 0, \tag{27}$$

must be satisfied at every point $P$ of the arc $C$ for all values $\zeta_i = \eta_i'$ belonging to an admissible arc through $P$. The condition (27) for the integral (24) reduces to the condition that

$$\sum_{ik} R_{ik}\zeta_i\zeta_k \geqq 0, \tag{28}$$

at every point of $C$, and equations $(26_2)$ become

$$\sum_k \psi_{\beta k}\zeta_k = 0, \tag{29}$$

where the arguments of $R_{ik}$ and $\psi_{\beta k}$ are the values $x, y_i(x), y_i'(x), \lambda_\beta(x)$

defining the arc $E_{12}$. We have therefore the following analogue of the Legendre condition, due to Clebsch:

*If the arc $E_{12}$ minimizes the integral* (1) *the condition* (28) *must be satisfied at every point of $E_{12}$, for all values $\zeta_i$ which satisfy equations* (29).

## 4. The Jacobi condition

In deriving the Jacobi condition by the methods of this paper the following lemma is useful:

Lemma 1. *Let $(u; \rho)$ be a solution of equations* (26) *such that the functions $u_i, u_i', \rho_\beta$, all vanish for some value $\xi$, $x_1 \leqq \xi \leqq x_2$. If the determinant*

$$R(x, y, y', \lambda) = \begin{vmatrix} R_{i1} \cdots R_{in}, & \psi_{1i} \cdots \psi_{mi} \\ \psi_{\beta 1} \cdots \psi_{\beta n}, & 0 \cdots 0 \end{vmatrix} \tag{30}$$

*is different from zero along the arc $E_{12}$, then*

$$u_i(x) \equiv 0, \qquad u_i'(x) \equiv 0, \qquad \rho_\beta(x) \equiv 0, \qquad x_1 \leqq x \leqq x_2.$$

Differentiate equations $(26_2)$ with respect to $x$. If $R \neq 0$ the resulting equations may then be solved with $(26_1)$ for $\eta_i''$ and $\mu_\beta'$. A system is thus obtained of the form

$$\frac{d\eta_i}{dx} = \eta_i', \qquad \frac{d\eta_i'}{dx} = g_i(x, \eta, \eta', \mu), \qquad \frac{d\mu_\beta}{dx} = h_\beta(x, \eta, \eta', \mu) \tag{31}$$

with second members linear and homogeneous in the variables $\eta_i$, $\eta_i'$, $\mu_\beta$. Every solution of (26) is also a solution of (31). Now it is well known* from the theory of differential equations that there is one, and but one, solution of equations (31) which assumes for a given value of $x$ a prescribed set of initial values. In particular, there is an unique solution which assumes the values $(0, 0, 0)$ for $x = \xi$. Since this solution is evidently

$$\eta_i \equiv 0, \qquad \eta_i' \equiv 0, \qquad \mu_\beta \equiv 0,$$

the lemma follows at once.

Assume now that a solution $(u\,; \rho)$ of equations (26) can be found in which the functions $u_i(x)$ vanish for two values $\xi_1, \xi_2$,

$$u_i(\xi_1) = 0 = u_i(\xi_2), \qquad x \leqq \xi_1 < \xi_2 < x_2, \tag{32}$$

without all being identically zero throughout the interval $(\xi_1, \xi_2)$. The curve $C$ in the $x\eta$-space defined by the conditions

$$\begin{aligned} &\eta_i \equiv 0, && \mu_\beta \equiv 0, && x_1 \leqq x \leqq \xi_1, \\ &\eta_i = u_i(x), && \mu_\beta = \rho_\beta(x), && \xi_1 \leqq x \leqq \xi_2, \\ &\eta_i \equiv 0, && \mu_\beta \equiv 0, && \xi_2 \leqq x \leqq x_2, \end{aligned}$$

* Bolza, loc. cit., p. 171.

is an admissible curve for the $x\eta$-problem, and has corner points possibly at $\xi_1$ and $\xi_2$. If the curve $C$ minimizes the integral $J''(O)$ the corner point condition

$$\Omega_{\eta_i'}(\xi_2 - 0) = \Omega_{\eta_i'}(\xi_2 + 0) \tag{33}$$

must be satisfied. Furthermore

$$\Phi_\beta(x, \eta, \eta')|_{\xi_2} = 0. \tag{34}$$

Since the values $\eta_i(\xi_2)$ are all zero, while

$$\eta_i'(\xi_2 - 0) = u_i'(\xi_2), \qquad \eta_i'(\xi_2 + 0) = 0,$$

the conditions (33) and (34) may be written

$$\begin{aligned} \sum_k 2u_k' R_{ik} + \sum_\beta \rho_\beta \psi_{\beta i}|_{\xi_2} &= 0, \\ \sum_k 2u_k' \psi_{\beta k} &= 0. \end{aligned} \tag{35}$$

If now the determinant (30) is different from zero at $x = \xi_2$ then the only solution of equation (35) is

$$u_k'(\xi_2) = 0, \qquad \rho_\beta(\xi_2) = 0. \tag{36}$$

But if equations (32) and the cornerpoint conditions (36) are satisfied it follows from Lemma 1 that

$$u_i(x) \equiv 0, \qquad \rho_\beta(x) \equiv 0, \qquad x_1 \leqq x \leqq x_2,$$

which contradicts the hypothesis that the functions $u_i(x)$ do not all vanish identically. Therefore the corner point conditions (36) are not satisfied, and consequently the curve $C$ cannot minimize the integral $J''(O)$.

Consider next the value of the integral $J''(O)$ taken over the curve $C$. Since equations $(26_2)$ are satisfied at every point of $C$, $\omega$ may be replaced by $\Omega$ and the integral $J''(O)$ for the $x\eta$-curve $C$ expressed in the form

$$J''(O) = \int_{\xi_1}^{\xi_2} \Omega(x, u, u', \rho)\, dx. \tag{37}$$

Now by Euler's theorem on homogeneous functions

$$2\Omega = \sum_k (u_k \Omega_{u_k} + u_k' \Omega_{u_k'}) + \sum_\beta \rho_\beta \Omega_{\rho_\beta},$$

where the last sum vanishes, since

$$\Omega_{\rho_\beta} = \Phi_\beta = 0$$

along $C$. The integral (37) then becomes

$$J''(O) = \frac{1}{2}\int_{\xi_1}^{\xi_2} \sum_k (u_k \Omega_{u_k} + u_k' \Omega_{u_k'})\, dx.$$

After integrating by parts, and using equations (25), it follows that

$$J''(O) = \tfrac{1}{2} \sum_k u_k \Omega_{u_k'} \Big]_{\xi_1}^{\xi_2} = 0,$$

on account of the conditions (32).

Since the curve $C$ does not minimize the integral $J''(O)$, and yet gives it the value zero, it follows that there must exist an admissible arc in the $x\eta$-space for which $J''(O)$ is negative and consequently the condition (24) not satisfied. In this event the arc $E_{12}$ does not minimize the integral (1). We may therefore state the following analogue of the Jacobi condition:

*Consider an arc $E_{12}$ that satisfies the conditions (A), . . ., (D) of section 1, and along which the determinant (30) is different from zero. Let $\xi_1$, $\xi_2$ be any two points*

$$x_1 \leqq \xi_1 \leqq x_2, \qquad x_1 < \xi_2 < x_2,$$

*where the values $x_1$, $x_2$ define the end points of the arc $E_{12}$. If $E_{12}$ minimizes the integral (1) there can exist no solution $(u\,;\rho)$ of equations (26) with elements $u_i(x)$ all vanishing at the end points of the interval $(\xi_1, \xi_2)$ but not identically zero within it.*

Two points $\xi_1$, $\xi_2$ with which there is associated a solution $(u\,;\rho)$ of equations (26) in which the functions $u_i(x)$ vanish at $\xi_1$ and $\xi_2$ without all being identically zero in $(\xi_1, \xi_2)$, are said to be *conjugate*. If, in particular $x_1 = \xi_1$ the Jacobi condition may be stated:

*If the arc $E_{12}$ minimizes the integral (1) no point $\xi$ conjugate to $x_1$ can lie within the interval $(x_1, x_2)$.*

## 5. Relations between the conjugate point and solutions of the Euler equations

In the literature of the calculus of variations the conjugate points are usually defined as the zeros of a determinant, the elements of which are obtained by differentiating the general solution of equations (2) and (10) with respect to the constants of integration. This definition furnishes methods of great practical value for determining conjugate points, and it is therefore desirable to show that these criteria may be obtained from the definition of conjugate points given in the preceding section. For this purpose it is convenient to prove the following lemmas.

Lemma 2. *If $(z_{ik}\,;r_{\beta k})$ is a system of $2n$ linearly independent solutions of equations (26) every other solution $(u\,;\rho)$ of these equations is expressible in the form*

$$u_i = \sum_k c_k z_{ik}, \qquad \rho_\beta = \sum_k c_k r_{\beta k}.$$

In the first place, the solutions of the system (31) satisfy the Jacobi equations $(26_1)$ and the conditions

$$\sum_i (\phi_{\beta i}\,\eta_i + \psi_{\beta i}\,\eta_i') = C_\beta. \tag{38}$$

By choosing the initial values of a solution of equations (31) properly it is always possible to find a solution $(z\,;\,r)$ of the Jacobi equations $(26_1)$ in which the functions $z_i(x)$ satisfy equations (38) for a prescribed set of values $C_\beta$. Consequently we can adjoin to the $2n$ solutions $(z_{ih}\,;\,r_{\beta h})$ of the lemma $m$ other solutions $(z_{i,\,2n+\alpha};\ r_{\beta,\,2n+\alpha})$ of the Jacobi equations, in which the functions $z_{i,\,2n+\alpha}$ satisfy the conditions

$$\sum_i (\phi_{\beta i}\,z_{i,\,2n+\alpha} + \psi_{\beta i}\,z_{i,\,2n+\alpha}') = \delta_{\beta\alpha}. \tag{39}$$

The $2n + m$ solutions of equations (31) thus obtained form a fundamental system. To prove this, write the product of the determinants $|\psi_{\beta\alpha}|$ and $D_1$ where

$$|\psi_{\beta\alpha}| = \begin{vmatrix} \psi_{\beta\alpha} & \psi_{\beta r} & \phi_{\beta i} & 0 \\ 0 & \delta_{sr} & 0 & 0 \\ 0 & 0 & \delta_{ji} & 0 \\ 0 & 0 & 0 & \delta_{\alpha\beta} \end{vmatrix}, \qquad D_1 = \begin{vmatrix} z_{i\nu}' \\ z_{i\nu} \\ r_{\alpha\nu} \end{vmatrix}.$$

In these determinants each element is a matrix whose dimensions are indicated by the subscripts attached, or else by its position. The ranges of the subscripts are those given at the beginning of section 1. The symbol $\delta_{ij}$ stands as usual for unity or zero according as $i = j$ or $i \neq j$.

Using the rows of $|\psi_{\beta\alpha}|$ and the columns of $D_1$, and remembering that the elements of the first $2n$ columns of $D_1$ are solutions of equations (26), while the elements of the last $m$ columns satisfy equations (39), we obtain the following product determinant of order $2n$:

$$D_2 = \begin{vmatrix} z_{rh}' \\ z_{ih} \\ r_{\alpha h} \end{vmatrix}.$$

Suppose now that $D_2$ vanishes for some value $x_3$, $x_1 \leqq x_3 \leqq x_2$. Then it is always possible to find $2n$ constants $c_h$ such that

$$\sum_h c_h z_{m+r,\,h}'(x_3) = 0, \qquad \sum_h c_h z_{ih}(x_3) = 0, \qquad \sum_h c_h r_{\beta h}(x_3) = 0. \tag{40}$$

Let

$$v_i(x) = \sum_h c_h\,z_{ih}(x), \qquad \sigma_\beta(x) = \sum_h c_h\,r_{\beta h}(x).$$

The functions $v_i$ are solutions of equations $(26_2)$. After substituting and putting $x = x_3$ it follows from (40) that

$$\sum_{\alpha} \psi_{\beta\alpha} v'_{\alpha}|_{x_3} = 0,$$

and consequently, since $|\psi_{\beta\alpha}| \neq 0$, that all the values $v'_{\alpha}(x_3)$ are zero. Since the functions $v_i, v'_i, \sigma_\beta$, all vanish at $x_3$ it follows from Lemma 1 that $v_i$ and $\sigma_\beta$ are identically zero. But this contradicts the hypothesis that the $2n$ solutions $(z_{ih}; r_{\beta h})$ are linearly independent. Therefore the determinants $D_2$ and $D_1$ cannot vanish anywhere in the interval $(x_1, x_2)$, and consequently the $2n + m$ solutions $(z_{iv}; r_{\beta v})$ of equations (31) form a fundamental system.

Let $(u; \rho)$ be any solution of equations (26). Then it is also a solution of equations (31) and is expressible in the form

$$(41) \qquad u'_i = \sum_{\nu} c_\nu z'_{i\nu}, \qquad u_i = \sum_{\nu} c_\nu z_{i\nu}, \qquad \rho_\beta = \sum_{\nu} c_\nu r_{\beta\nu}.$$

Multiply equations $(41_1)$ and $(41_2)$ by $\psi_{\beta i}$ and $\phi_{\beta i}$ respectively and add the resulting $2n$ equations. Since the functions $u_i$ and $z_{ih}$ are solutions of equations (26), while the functions $z_{i,\,2n+\alpha}$ satisfy equations (39), it follows that the constants $c_{2n+\alpha}$ are all zero. Therefore

$$u_i = \sum_{h} c_h z_{ih}, \qquad \rho_\beta = \sum_{h} c_h r_{\beta h},$$

which proves the lemma.

LEMMA 3. *If $(z_{ih}; r_{\beta h})$ is a system of $2n$ solutions of equations* (26), *and if the determinant*

$$\Delta(x, x_1) = \begin{vmatrix} z_{ih}(x) \\ z_{ih}(x_1) \end{vmatrix}$$

*does not vanish identically, then the zeros of $\Delta(x, x_1)$ determine the conjugate points to* 1.

First, let 3 be a point conjugate to 1. Then there is associated with the point 3 a solution $(u; \rho)$ of equations (26) in which the functions $u_i(x)$ vanish at $x_1$ and $x_3$ without all being identically zero throughout the interval $(x_1, x_3)$. Since $\Delta(x, x_1) \not\equiv 0$ the $2n$ solutions $(z_{ih}; r_{\beta h})$ are linearly independent, and consequently by Lemma 2 the functions $u_i(x)$ are expressible in the form

$$u_i(x) = \sum_{h} c_h z_{ih}.$$

If now the functions $u_i(x)$ all vanish at $x_1$ and $x_3$ the conditions

$$(42) \qquad \sum_{h} c_h z_{ih}(x_1) = 0, \qquad \sum_{h} c_h z_{ih}(x_3) = 0,$$

must be satisfied. This requires that $\Delta(x_3, x_1) = 0$. Therefore the abscissa $x_3$ of every point 3 conjugate to 1 is a zero of $\Delta(x, x_1)$.

Next, let $\Delta(x_3, x_1) = 0$. Then it is always possible to determine $2n$ constants $c_h$ such that equations (42) are satisfied. If we define

$$u_i = \sum_k c_k z_{ik}, \qquad \rho_\beta = \sum_k c_k r_{\beta k},$$

we have a solution of equations (26) in which the functions $u_i(x)$ vanish at $x_1$ and $x_3$ and yet are not all identically zero in $(x_1, x_3)$ on account of the hypothesis that $\Delta(x, x_1) \not\equiv 0$. Therefore every zero $x_3$ of the determinant $\Delta(x, x_1)$ is the abscissa of a point 3 which is conjugate to 1.

LEMMA 4. *If $(z_{in}\,;\, r_{\beta n})$ is a system of $n$ linearly independent solutions of equations* (26) *such that for all values of $i$ and $k$*

$$z_{ik}(x_1) = 0, \tag{43}$$

*then every other solution $(u\,;\,\rho)$ in which the functions $u_i(x)$ all vanish for $x = x_1$ is expressible in the form*

$$u_i = \sum_k c_k z_{ik}, \qquad \rho_\beta = \sum_k c_k r_{\beta k}.$$

The proof is quite similar to that of Lemma 2. In the present instance we adjoin to the $n$ solutions $(z_{ik}\,;\, r_{\beta k})$ of the lemma $n$ solutions $(z_{i,\,n+k}\,;\, r_{\beta,\,n+k})$ of equations (26), with the single hypothesis that

$$|z_{i,\,n+k}(x_1)| \neq 0, \tag{44}$$

and in addition the $m$ solutions $(z_{i,\,2n+a}\,;\, r_{\beta,\,2n+a})$ of equations (31) which were used in the proof of Lemma 2. The $2n + m$ solutions of equations (31) thus obtained form a fundamental system. To show this form the product determinant $D_2$ and assume that it vanishes for some value $x_3$, $x_1 \leqq x_3 \leqq x_2$. By precisely the same argument as that used in Lemma 2 it follows that

$$\sum_k c_k z_{ik}(x) \equiv 0, \qquad \sum_k c_k r_{\beta k}(x) \equiv 0. \tag{45}$$

On putting $x = x_1$ in equations (45) and using (43) and (44) it will be seen that all the $n$ constants $c_{n+k}$ are zero. Equations (45) then become

$$\sum_k c_k z_{ik}(x) \equiv 0, \qquad \sum_k c_k r_{\beta k}(x) \equiv 0,$$

which contradicts the hypothesis that the $n$ solutions $(z_{ik}\,;\, r_{\beta k})$ are linearly independent. Therefore the determinants $D_2$ and $D_1$ cannot vanish anywhere in $(x_1, x_2)$, and consequently the $2n + m$ solutions $(z_{iv}\,;\, r_{\beta v})$ of equations (31) form a fundamental system. It follows at once from Lemma 2 that every solution $(u\,;\,\rho)$ of equations (26) is expressible in the form

$$u_i = \sum_k c_k z_{ik}, \qquad \rho_\beta = \sum_k c_k r_{\beta k}. \tag{46}$$

If the functions $u_i(x)$ all vanish at $x_1$ then

$$\sum_k c_k z_{ik}(x_1) = 0,$$

and from the relations (43) and (44) it follows that the $n$ constants $c_{n+k}$ are all zero. Equations (46) then become

$$u_i = \sum_k c_k z_{ik}, \qquad \rho_\beta = \sum_k c_k r_{\beta k}.$$

LEMMA 5. *If $(z_{ik}; r_{\beta k})$ is a system of $n$ solutions of equations* (26) *in which the functions $z_{ik}(x)$ all vanish at $x_1$, and if the determinant*

$$D(x) = |z_{ik}|$$

*does not vanish identically, then the zeros of $D(x)$ determine the conjugate points to* 1.

First, let 3 be a point conjugate to 1, and $(u; \rho)$ the associated solution. By Lemma 4

$$u_i = \sum_k c_k z_{ik}.$$

Since $u_i(x_3) = 0$ the conditions

$$\sum_k c_k z_{ik}(x_3) = 0$$

must be satisfied, and therefore $D(x_3) = 0$, that is, the abscissa $x_3$ of every point 3 conjugate to 1 is a zero of $D(x)$. Again, suppose that

$$D(x_3) = 0.$$

In this event $n$ constants $c_k$ can be determined so that

$$\sum_k c_k z_{ik}(x_3) = 0.$$

If we define

$$u_i = \sum_k c_k z_{ik}, \qquad \rho_\beta = \sum_k c_k r_{\beta k},$$

we have a solution of equations (26) in which the functions $u_i(x)$ vanish at $x_1$ and $x_3$ and yet are not all identically zero, on account of the hypothesis $D(x) \not\equiv 0$. Therefore every zero $x_3$ of $D(x)$ is the abscissa of a point 3 which is conjugate to 1.

As consequences of the foregoing lemmas we have the customary criteria for determining the conjugate points, as follows:

THEOREM. *Let*

$$y_i = y_i(x, c_1, \cdots, c_{2n}), \qquad \lambda_\beta = \lambda_\beta(x, c_1, \cdots, c_{2n}) \tag{47}$$

*be a $2n$ parameter family of solutions of the Euler equations* (10) *and the equations of condition* (2), *containing the minimizing arc $E_{12}$ and its multipliers for the values $(c_1, \cdots, c_{2n}) = (0, \cdots, 0)$. If the determinant*

$$\Delta(x, x_1) = \begin{vmatrix} \dfrac{\partial y_i(x)}{\partial c_k} \\ \dfrac{\partial y_i(x_1)}{\partial c_k} \end{vmatrix}$$

*does not vanish identically along* $E_{12}$, i. e., *when the constants* $c_h$ *are all zero, then the zeros of* $\Delta(x, x_1)$ *determine the conjugate points to* 1 *on* $E_{12}$.

If the functions (47) are substituted in equations (2) and (10) identities in $x, c_1, \cdots, c_{2n}$ are obtained.* Differentiating with respect to any parameter $c_h$ and then setting $(c_1, \cdots, c_{2n}) = (0, \cdots, 0)$ we obtain the well-known result that each of the $2n$ systems of functions

$$\frac{\partial y_i}{\partial c_h}, \quad \frac{\partial \lambda_\beta}{\partial c_h} \qquad (h = 1, 2, \cdots, 2n),$$

in which the constants $c_h$ are all zero, is a solution of equations (26) formed for the extremal arc $E_{12}$. Since the linear independence of the $2n$ solutions is a consequence of the hypothesis that $\Delta(x, x_1) \not\equiv 0$ the theorem follows at once from Lemma 3.

THEOREM. *Let*

$$y_i = y_i(x, c_1, \cdots, c_n), \qquad \lambda_\beta = \lambda_\beta(x, c_1, \cdots, c_n)$$

*be an* $n$ *parameter family of solutions of equations* (2) *and* (10), *containing the minimizing arc* $E_{12}$ *for the values* $(c_1, \cdots, c_n) = (0, \cdots, 0)$, *and satisfying the conditions*

$$y_i(x_1, c_1, \cdots, c_n) = y_i(x_1). \tag{48}$$

*If the determinant*

$$D(x) = \left| \frac{\partial y_i(x)}{\partial c_k} \right|$$

*does not vanish identically along* $E_{12}$ *then the zeros of* $D(x)$ *determine the conjugate points to* 1 *on* $E_{12}$.

The $n$ systems of functions

$$\frac{\partial y_i}{\partial c_k}, \quad \frac{\partial \lambda_\beta}{\partial c_k} \qquad (k = 1, 2, \cdots, n),$$

where $(c_1, \cdots, c_n) = (0, \cdots, 0)$, are solutions of equations (26). The linear independence of these solutions follows from the hypothesis that $D(x) \not\equiv 0$. Furthermore, it is clear from equations (48) that

$$\frac{\partial y_i(x_1)}{\partial c_k} = 0.$$

The theorem is an immediate consequence of Lemma 5.

* In this connection see Bolza, loc. cit., pp. 72, 623.

# ON TRANSCENDENTAL NUMBERS*

BY

AUBREY J. KEMPNER

In 1851, Liouville† gave the following theorem (the proof of which is very simple):

Let $f(z) = 0$ be an equation of degree $n \geqq 2$ with real integral coefficients and irreducible in the domain $R(1)$, $z$ a real root of this equation, and $p/q$ any real rational number. Then a positive number $A$ can be found, which is independent of $n$, such that for all $p/q$,

$$\left| z - \frac{p}{q} \right| > \frac{1}{Aq^n}.$$

By applying his theorem to numbers of the form $\sum_{\nu=1}^{\infty} (\alpha_\nu / l^{\nu!})$, $l$ a positive integer $\geqq 2$, $\alpha_\nu$ any integer whose absolute value is limited, Liouville showed that all such numbers are transcendental.

By choosing $l = 10$, a rule is obtained for representing in decimal-fraction form the numbers contained in a certain non-enumerable set of transcendental numbers.‡

From Liouville's method of proof it follows that all numbers $\sum_{\nu=0}^{\infty} (\alpha_\nu / l^{\gamma_\nu})$, when only $\gamma_\nu$ increases "sufficiently rapidly," are transcendental.

E. Maillet§ proves in his book the transcendency of the members of certain sets of numbers, the simplest of which are represented by series of the type

$$\sum_{n=0}^{\infty} (a_n \div b^{nb^n}) \cdot x^n$$

for all rational and even for all algebraic $x$, and G. Faber‖ uses a generalization of Liouville's theorem and treats a more general type of series which may

---

* Presented to the Society, December 30, 1915.

† Journal de mathématiques pures et appliquées, vol. 16 (1851), p. 133.

‡ The first proof of the existence of transcendental numbers was also given by Liouville, in the Comptes-rendus, vol. 18 (1844), p. 883, p. 910 (reproduced in †), and was based on an investigation of continued fractions.

§ *Introduction à la théorie des nombres transcendants et des propriétés arithmétiques des fonctions*, Paris (1906), p. 105.

‖ Mathematische Annalen, vol. 58 (1904), p. 545.

by this method be shown to yield transcendental numbers, namely all series of the form

$$f(x) = \sum_{n=0}^{\infty} \frac{h_n \cdot x^n}{(p_1 \cdot p_2 \cdots p_n)^{q_1 \cdot q_2 \cdots q_n}},$$

where the $h_n$, $p_n$, $q_n$ are integers, $h_n$ finite (or growing infinite with $n$ in some particular fashion),

$$p_n > 1, \qquad \lim_{n=\infty} q_n = \infty;$$

and he shows that $f(x)$ is a transcendental number for all algebraic $x$. These seem to be about the only advances along this line.* Series of the type $\sum_{n=0}^{\infty}(\alpha_n/a^{b^n})$ can not be treated by these methods.

The object of the present paper is to prove the following

THEOREM. *Let $a$ be an integer greater than 1, $p/q$ a rational fraction, $p \gtrless 0$, $q > 0$; $\alpha_n$, $(n = 0, 1, 2, \cdots)$, any positive or negative integer smaller in absolute value than a fixed arbitrary number $M$, but only a finite number of the $\alpha_n$ equal to 0; then $\sum_{n=0}^{\infty}(\alpha_n/a^{2^n}) \cdot (p/q)^n$ is always a transcendental number.*

We shall prove the theorem in the form just stated. However, the proof still holds, with only slight modifications, for either or both of the following extensions:

1. The exponent $2^n$ may be replaced everywhere by $b^n$, $b$ any fixed positive integer greater than 2.

2. $\alpha_n$ need not be limited; for example, the number is transcendental when $|\alpha_n| < R^n$, $R$ an arbitrary fixed number; or $\alpha_n$ may be a rational number $\delta/\epsilon$, $|\delta| < R^n$, $\epsilon < R^n$, but always with the restriction that only a finite number of the $\alpha_n$ equal to 0.

*Proof.* All letters to be used will denote real numbers and all are integers except $z$ and $c_3$. The symbol $r^{s^t}$ means $r^{(s^t)}$, $r^{s^{t^u}}$, means $r^{(s^{(t^u)})}$, etc.

Max. $(a_1, a_2, a_3, \cdots, a_\tau)$ denotes the largest of the positive values $a_1, a_2, a_3, \cdots, a_\tau$.

Let $z = \sum_{n=0}^{\infty}(\alpha_n/a^{2^n}) \cdot (p/q)^n$, with the restrictions on the $\alpha_n$, $a$, $p$, and $q$ mentioned in the theorem.

If $z$ is not a transcendental number, there must exist a certain equation

$$f(z) = \sum_{\mu=0}^{k} A_\mu z^\mu = 0,$$

where $k$ is a fixed positive integer and the $A_\mu$ are integers, $A_k \neq 0$. Let $N$ be a positive integer, such that $N > |A_\mu|$, $(\mu = 0, 1, \cdots, k)$. We shall

* E. Borel, *Leçons sur la théorie de la croissance, recueillies et redigées par* A. Denjoy, Paris (1910), Chapter V, and Axel Thue, Journal für Mathematik, vol. 135 (1909), p. 284, particularly the latter, have generalized Liouville's theorem in important respects, but these generalizations have so far not had any influence on the investigation of transcendental numbers.

substitute $z = \sum_{n=0}^{\infty} (\alpha_n/a^{2^n}) \cdot (p/q)^n$ in $f(z)$ and shall show that $f(z) \neq 0$ for any given $f(z)$.

Once chosen, all of the following numbers are to be considered constant: $\alpha_n$ $(n = 0, 1, \cdots)$; $a$; $p$; $q$; $k$; $A_\mu$ $(\mu = 0, 1, 2, \cdots, k)$; $M$; $N$.

Since $\sum_{n=0}^{\infty} (\alpha_n/a^{2^n}) \cdot x^n$ is a power-series, convergent for all values of $x$, we may substitute $z$ in $f(z)$ and rearrange terms as we like. After substituting, we shall arrange the terms according to increasing denominators, *without canceling anywhere,* and collect terms with equal denominators.

The denominators are formed by taking the product of $k$ or fewer factors of the form $a^{2^\nu} \cdot q^\nu$, repetition admitted.

We prove first the following statement:

*When at the same time $n > k$ and $n > n_1$, where $n_1$ is a certain positive integer which will be characterized in the proof, then the three numbers*

$$\gamma_1 = a^{2^{n-1}+2^{n-2}+\cdots+2^{n-k+1}+2^{n-k-1}} \cdot q^{(n-1)+(n-2)+\cdots+(n-k+1)+(n-k-1)},$$

$$\gamma_2 = a^{2^{n-1}+2^{n-2}+\cdots+2^{n-k+1}+2^{n-k}} \cdot q^{(n-1)+(n-2)+\cdots+(n-k+1)+(n-k)},$$

$$\gamma_3 = a^{2^n} \cdot q^n$$

*satisfy the two conditions*:

(1) $$\gamma_1 < \gamma_2 < \gamma_3,$$

(2) $$\gamma_1, \gamma_2, \gamma_3$$

*are three consecutive denominators of our fractions in* $\sum_{\mu=0}^{k} A_\mu z^\mu$.

It is clear that all denominators containing any factor $a^{2^\nu}$, $\nu > n$, are larger than $\gamma_3$, and of all denominators containing the factor $a^{2^n}$, $\gamma_3$ is the smallest. Consequently all denominators smaller than $\gamma_3$ contain only factors $a^{2^\nu} \cdot q^\nu$, $(\nu = 0, 1, \cdots, n-1)$. Of all denominators smaller than $\gamma_3$ containing not more than $k$ of these factors, $(n > k)$, it is obvious that there is none between $\gamma_2$ and $\gamma_1$ and none between $\gamma_2$ and $\gamma_3$. It remains however to be shown that $\gamma_2 < \gamma_3$, in spite of the higher powers of $q$ involved in $\gamma_2$, that is:

$$a^{2^{n-1}+2^{n-2}+\cdots+2^{n-k}} \cdot q^{(n-1)+(n-2)+\cdots+(n-k)} < a^{2^n} \cdot q^n$$

for all $n$ from a certain value on. We have

$$q^{(n-1)+(n-2)+\cdots+(n-k)-n} < a^{2^{n-k}},$$

which is *a fortiori* satisfied when

$$q^{n^2/2} = (q^{\frac{1}{2}})^{n^2} < a^{2^{n-k}} = (a^{2^{-k}})^{2^n};$$

this is true from a certain integer $n = n_1$ on. Thus our statement is proved, and we shall henceforth take $n$ greater than Max $(k, n_1)$.

To prove that $f(z) \neq 0$, we shall show that $f(z)$ may be written in the form

$$f(z) = \frac{c_1}{\gamma_1} + \frac{c_2}{\gamma_2} + c_3,$$

satisfying the following conditions:

(*a*) $c_1, c_2$ are integers, $c_1 \gtreqless 0, c_2 \neq 0$; $c_3$ a real number $\gtreqless 0$; $\gamma_1, \gamma_2$ numbers of the type defined above, so that $\gamma_2$ is a multiple of $\gamma_1$; $\gamma_2 = l \cdot \gamma_1, l > 1$.

(*b*) $\dfrac{|c_2|}{\gamma_2} < \dfrac{1}{\gamma_1}$.

(*c*) $|c_3| < \dfrac{1}{\gamma_2}$.

We must admit four possible cases:

If $c_1 = 0, c_3 = 0$: $f(z) \neq 0$ because $c_2 \neq 0$.

If $c_1 = 0, c_3 \neq 0$: $f(z) = \dfrac{c_2}{\gamma_2} + c_3 \neq 0$ from (*a*) and (*c*), without using (*b*).

If $c_1 \neq 0, c_3 = 0$: $f(z) = \dfrac{c_1}{\gamma_1} + \dfrac{c_2}{\gamma_2} \neq 0$ from (*a*) and (*b*), without using (*c*).

If $c_1 \neq 0, c_3 \neq 0$: $\dfrac{|c_1|}{\gamma_1} = \dfrac{l|c_1|}{\gamma_2} > \dfrac{c_2}{\gamma_2}$, therefore

$$\frac{|c_1|}{\gamma_1} - \frac{|c_2|}{\gamma_2} \geqq \frac{1}{\gamma_2} > |c_3|, \qquad \frac{|c_1|}{\gamma_1} > \frac{|c_2|}{\gamma_2} + |c_3|,$$

and hence

$$f(z) = \frac{c_1}{\gamma_1} + \frac{c_2}{\gamma_2} + c_3 \neq 0.$$

*Always assuming that no cancellations have been performed, either in*

$$\sum_{n=0}^{\infty} (\alpha_n / a^{2^n}) \cdot (p/q)^n \quad \text{or in} \quad \sum_{\mu=0}^{k} A_\mu z^\mu,$$

our $c_2/\gamma_2$ shall consist of all terms which have exactly the denominator

$$\gamma_2 = a^{2^{n-1}+2^{n-2}+\cdots+2^{n-k}} \cdot q^{(n-1)+(n-2)+\cdots+(n-k)},$$

$n$ being properly chosen. Evidently $c_2/\gamma_2$ arises entirely from the term $A_k z^k$ of $f(z)$, and we obtain

$$\frac{c_2}{\gamma_2} = \frac{A_k \cdot \alpha_{n-1} \cdot \alpha_{n-2} \cdots \alpha_{n-k} \cdot p^{(n-1)+(n-2)+\cdots+(n-k)}}{a^{2^{n-1}+2^{n-2}+\cdots+2^{n-k}} \cdot q^{(n-1)+(n-2)+\cdots+(n-k)}} \cdot \frac{k!}{1!\,1!\cdots 1!},$$

$$\frac{|c_2|}{\gamma_2} \leqq \frac{k!\, N \cdot M^k \cdot |p|^{(n-1)+(n-2)+\cdots+(n-k)}}{a^{2^{n-1}+2^{n-2}+\cdots+2^{n-k}} \cdot q^{(n-1)+(n-2)+\cdots+(n-k)}},$$

and $c_2 \neq 0$, provided $n > n_2$, where $n_2$ is so large that $\alpha_{n_2-k}$ and all following coefficients are different from zero.

Our $c_1/\gamma_1$ shall comprise all terms with denominators smaller than $\gamma_2$. If $c_1 = 0$, $c_1/\gamma_1$ is not needed. If $c_1 \neq 0$, then

$$\gamma_1 = a^{2^{n-1}+2^{n-2}+\cdots+2^{n-k+1}+2^{n-k-1}} \cdot q^{(n-1)+(n-2)+\cdots+(n-k+1)+(n-k-1)}$$

is common denominator of all fractions under consideration, which can therefore be added together, giving $c_1/\gamma_1$.

To prove $|c_2|/\gamma_2 < 1/\gamma_1$, we show that, from a certain $n = n_3$ on,

$$\frac{k!\,N \cdot M^k \cdot |p|^{(n-1)+(n-2)+\cdots+(n-k)}}{a^{2^{n-1}+2^{n-2}+\cdots+2^{n-k}} \cdot q^{(n-1)+(n-2)+\cdots+(n-k)}} < \frac{1}{a^{2^{n-1}+2^{n-2}+\cdots+2^{n-k+1}+2^{n-k-1}} \cdot q^{(n-1)+(n-2)+\cdots+(n-k+1)+(n-k-1)}}.$$

We assume $n > \text{Max}\,(k, n_1, n_2)$. Our inequality reduces to

$$\frac{k!\,N \cdot M^k}{q} \cdot |p|^{(n-1)+(n-2)+\cdots+(n-k)} < a^{2^{n-k-1}},$$

which is *a fortiori* satisfied if

$$C_1 \cdot |p|^{n^2/2} < a^{2^{n-k-1}},$$

where

$$C_1 = \frac{k!\,N \cdot M^k}{q}$$

is independent of $n$;

$$C_1 \cdot |p_\nu^{\frac{1}{2}}|^{n^2} < (a^{2^{-k-1}})^{2^n},$$

which is true from a certain $n_3$ on. We assume $n > \text{Max}\,(k, n_1, n_2, n_3)$. This proves $(b)$.

Our $c_3$ shall consist of all terms with denominators greater than $\gamma_2$, that is, of all terms of $f(z)$ not accounted for by $c_1/\gamma_1 + c_2/\gamma_2$. Hence $c_3$ is built up of fractions of the form

$$\frac{A_\mu \cdot \alpha_{\tau_1} \cdot \alpha_{\tau_2} \cdots \alpha_{\tau_\rho} \cdot p^{\tau_1+\tau_2+\cdots+\tau_\rho}}{a^{2^{\tau_1}+2^{\tau_2}+\cdots+2^{\tau_\rho}} \cdot q^{\tau_1+\tau_2+\cdots+\tau_\rho}},$$

where $\rho \leqq k$, and $\tau_1, \tau_2, \cdots, \tau_\rho$ are numbers of the sequence $0, 1, 2, \cdots$ not necessarily different from each other.

These denominators are all of the form $a^{\lambda_1} \cdot q^{\lambda_2}$, where the exponents of $a$ are formed by taking the sum of $k$ or less numbers of the infinite sequence $2^\nu$, $(\nu = 0, 1, 2, \cdots)$, repetition permitted, with the restriction that at least one of the $k$ or fewer numbers $2^\nu$ shall be greater than or equal to $2^n$, where $n > \text{Max}\,(k, n_1, n_2, n_3)$. There are, as is easily seen, less than $(n+2)^k$ exponents that can be so formed from the first $n+1$ numbers $2^0, 2^1, \cdots, 2^n$, counting exponents separately even when they differ only in the order of their

summands $2^{\tau_1}, 2^{\tau_2}, \cdots, 2^{\tau_\rho}$. In the same way it is seen that there are less than $(n+3)^k$ exponents formed by taking only numbers of the set $2^0, 2^1, \cdots, 2^{n+1}$; less than $(n+l+2)^k$ by taking only numbers of the set $2^0, 2^1, \cdots, 2^{n+l}$ for $l = 1, 2, 3, \cdots$. The denominators $a^{\lambda_1} \cdot q^{\lambda_2}$ all contain also factors $q^{\lambda_2}$. Hence there are certainly less than $(n+2)^k$ fractions with denominators smaller than $a^{2^n}$, less than $(n+3)^k$ fractions with denominators $d$, where $a^{2^n} \leqq d < a^{2^{n+1}}$, because there are altogether less than $(n+3)^k$ fractions with denominators less than or equal to $a^{2^{n+1}}$, and less than $(n+l+2)^k$ fractions with denominators $d$, where $a^{2^{n+l-1}} \leqq d < a^{2^{n+l}}$. I increase (or at least do not decrease) the absolute value of all fractions by replacing $q^{\lambda_2}$ by 1.

Those numerators belonging to denominators $d$, $a^{2^n} \leqq d < a^{2^{n+1}}$, have all of their $\tau_1, \tau_2, \cdots, \tau_\rho$ not larger than $n$, those belonging to denominators $d$, $a^{2^{n+1}} \leqq d < a^{2^{n+2}}$, have all of their $\tau \leqq n+1$, and those fractions with denominators $d$, $a^{2^{n+l}} \leqq d < a^{2^{n+l+1}}$, have all $\tau \leqq n+l$.

Altogether we find for $c_3$:

$$|c_3| < N \cdot M^k \cdot \left[ \frac{(n+2)^k \cdot |p|^{kn}}{a^{2^n}} + \frac{(n+3)^k \cdot |p|^{k(n+1)}}{a^{2^{n+1}}} + \cdots + \frac{(n+l+2)^k \cdot |p|^{k(n+l)}}{a^{2^{n+l}}} + \cdots \right].$$

The convergence of this expression is obvious. Besides taking $n$ greater than Max. $(k, n_1, n_2, n_3)$ we now take $n$ so large that the sum in brackets is smaller than

$$\frac{2(n+2)^k \cdot |p|^{kn}}{a^{2^n}}.$$

This is certainly true when the ratio of two consecutive terms is always smaller than $\frac{1}{2}$, and happens, for example, when $n > \log_2 [k \cdot \log_a |4p|]$, as is easily verified. Let $n_4$ be an integer satisfying this relation, and take $n$ greater than Max. $(k, n_1, n_2, n_3, n_4)$. Substituting, we have

$$|c_3| < \frac{2N \cdot M^k \cdot (n+2)^k \cdot |p|^{kn}}{a^{2^n}},$$

and we shall show that the expression on the right is, for sufficiently large values of $n$, less than

$$\frac{1}{a^{2^{n-1}+2^{n-2}+\cdots+2^{n-k}} \cdot q^{(n-1)+(n-2)+\cdots+(n-k)}},$$

thus establishing the inequality $|c_3| < 1/\gamma_2$ and proving our theorem.

Our inequality for $n$ reduces to

$$2N \cdot M^k \cdot (n+2)^k \cdot |p|^{kn} \cdot q^{(n-1)+(n-2)+\cdots+(n-k)} < a^{2^{n-k}},$$

which is *a fortiori* satisfied when

$$C_2 \cdot (n+2)^k \cdot |p^k|^n \cdot (q^k)^{n^2} < (a^{2^{-k}})^{2^n},$$

where $C_2 = 2N \cdot M^k$ is independent of $n$.

By making $n$ sufficiently large we can satisfy the following three inequalities simultaneously:

$$C_2 \cdot (n+2)^k < |p^k|^n, \qquad |p^{2k}|^n < (q^k)^{n^2}, \qquad q^{n^2} < (a^{2^{-k}})^{2^n},$$

which, combined, prove our inequality for all $n$ greater than a certain integer $n_6$. By taking $n$ greater than Max. $(k, n_1, n_2, n_3, n_4, n_5, n_6)$ we meet all restrictions which have been successively imposed on $n$ during the proof.

The condition that only a finite number of coefficients shall be zero (in order to ensure $c_2 \neq 0$), I have not been able to remove.

By taking $p/q = 1$, we see that all numbers $\sum_{n=0}^{\infty} (\alpha_n / a^{2^n})$, ($\alpha_n$ an integer, $|\alpha_n|$ limited, and from a certain point on $|\alpha_n| \geqq 1$) are transcendental.

As another special case we mention the function, $\sum_{n=0}^{\infty} x^{2^n}$, introduced by Fredholm* to demonstrate the existence of analytical functions possessing certain peculiar properties on their natural boundaries. It follows from our theorem that this function has transcendental values for an infinite set of real rational values of the argument having the origin as a limiting point.

* See Mittag-Leffler, Acta Mathematica, vol. 15 (1891), p. 279.

URBANA, ILLINOIS

# THE LINEAR DEPENDENCE OF FUNCTIONS OF SEVERAL VARIABLES, AND COMPLETELY INTEGRABLE SYSTEMS OF HOMOGENEOUS LINEAR PARTIAL DIFFERENTIAL EQUATIONS*

BY

GABRIEL M. GREEN

## INTRODUCTION

In the case of $n$ functions of a single variable, the vanishing of the wronskian is the most familiar criterion for their linear dependence. The wronskian also plays an important part in connection with the theory of a single ordinary homogeneous linear differential equation of the $n$th order, even in questions which do not concern the linear dependence of solutions of the equation. It is the purpose of the present paper to generalize the fundamental facts connected with these topics, to the case of functions of several variables.

The characteristic property of an ordinary homogeneous linear differential equation is that any solution is expressible linearly, with constant coefficients, in terms of a fundamental set of solutions. The natural generalization to several independent variables is afforded by the completely integrable system of homogeneous linear partial differential equations, in one dependent variable, any solution of which is likewise linearly dependent on a finite number of solutions of the system. It is such systems which are discussed in the latter part of this paper; the subject of the linear dependence of functions of several variables is developed in the first part, chiefly with a view to its subsequent application to completely integrable systems.

In our discussion of linear dependence, we consider throughout $n$ functions $y_1, y_2, \cdots, y_n$ of $p$ independent variables $u_1, u_2, \cdots, u_p$. Functions and variables may be either real or complex; if any of the variables be complex, we suppose the functions to be analytic in those variables. All of the main theorems, however, are stated and proved under the supposition that the independent variables are real; the modifications—usually simplifications and omissions—are easily supplied for cases in which one or more of the independent variables is complex. The $n$ functions, then, of the real variables

* Presented to the Society in two communications under different titles, October 31, 1914, and October 30, 1915. The results concerning the linear dependence of functions of several variables were read on the latter date.

$u_1, u_2, \cdots, u_p$, exist, together with such of their derivatives as come into the discussion, in a $p$-dimensional region, open or closed, which for simplicity may be taken to be a $p$-dimensional parallelopiped. This region we denote throughout by $A$. We shall suppose that all partial differentiations taken with respect to more than one of the independent variables are independent of the order in which the differentiations are effected. Thus, by the existence of $\partial^2 y/\partial u \partial v$ we imply the existence of $\partial^2 y/\partial v \partial u$ and the identity of these two derivatives.

In the case of functions of several variables, there is no one determinant which may properly be taken as the generalization of the wronskian. In fact, since there is no relation between the number of functions and the number of independent variables, no definite form can be prescribed for such a determinant; and moreover, the difficulty in this direction is increased by the fact that we do not assume our functions to be analytic. However, in the general criteria which we obtain concerning the linear dependence of $n$ functions, a matrix takes the place of the wronskian. This matrix contains $n$ columns, and each row is formed of the same derivative of the $n$ functions. Evidently, the identical vanishing of all the $n$-rowed determinants of such a matrix is a necessary condition for the linear dependence of the $n$ functions. The sufficient conditions at which we arrive are expressed in terms of the vanishing of the $n$-rowed determinants of a matrix of the kind just described, and are therefore also necessary—a point which we shall not mention in the body of the paper. The case of $n$ functions of a single variable is of course included in our general theory; the matrix for such functions may have $n$ or more rows, and if it has only $n$ the determinant of the matrix is precisely the wronskian of the $n$ functions.

So far as we are aware, the only completely integrable systems which have been studied are of two kinds. Of these, the only non-analytic ones are of the first order, and are in fact simply systems of total differential equations.* For systems consisting of equations of order higher than the first, an analytic character has been presupposed† in establishing the existence of solutions. By a mere change of notation, however, we cause such a system to take on the form of a system of total differential equations, so that the existence of solutions may be proved in the non-analytic case for completely integrable systems consisting of equations of unrestricted orders.

Although for a set of functions of several variables there is in general no distinctive determinant, formed from partial derivatives of the functions,

* Cf., for instance, Mayer, *Ueber unbeschränkt integrable Systeme von linearen totalen Differentialgleichungen und die simultane Integration linearer partieller Differentialgleichungen*, Mathematische Annalen, vol. 5 (1872), pp. 448–470.

† Cf. C. Riquier, *Les systèmes d'équations aux dérivées partielles*, Paris, Gauthier-Villars, 1910.

which might play the rôle of a wronskian, we have been able to define a wronskian for a fundamental set of solutions of a completely integrable system. By means of this wronskian, if we may so term it, we find certain properties of homogeneous linear completely integrable systems quite analogous to those of an ordinary differential equation, linear and homogeneous, of the $n$th order. An important result in this connection is a proof that a large and important class of completely integrable systems of the kind considered may be put into a so-called canonical form by a certain transformation of the dependent variable. The theorem is analogous to the familiar one in the case of an ordinary differential equation

$$\frac{d^n y}{dx^n} + p_1 \frac{d^{n-1} y}{dx^{n-1}} + p_2 \frac{d^{n-2} y}{dx^{n-2}} + \cdots + p_n y = 0,$$

according to which the equation can be reduced in essentially but one way to a unique form for which the coefficient $p_1$ is zero, by a transformation of the dependent variable alone of the form $\bar{y} = \lambda(x) y$.

The reduction, just referred to, of a completely integrable system to a canonical form, is very useful in certain geometric questions. Completely integrable systems have been important in differential geometry, but not until recently has systematic use been made of them in connection with that branch of the subject for which they are most peculiarly fitted, viz., projective differential geometry. We owe to Wilczynski a very general method for dealing with questions in that field; in fact, according to this method, the projective differential geometry of a configuration is equivalent to the theory of the invariants and covariants, under certain continuous groups, of a completely integrable system, the configuration being defined by a fundamental set of solutions of the system. The continuous groups which give rise to the invariants and covariants are generally broken up into two parts: a transformation of the independent variables, and a transformation of the dependent variable. The coefficients of the canonical form are invariant under this latter transformation, and in terms of these coefficients and their derivatives may be expressed all quantities having the same invariantive property. It is therefore highly desirable to have a method for computing these seminvariants, as they are called; we show that *this can be done in a purely mechanical way for certain completely integrable systems.* In the various particular problems studied by Wilczynski and others, this method is the one which has always been followed.

As an illustration, we show in § 7 how the foregoing may be applied to the theory of curvilinear coördinates in space of $n$ dimensions.

## 1. The fundamental theorem

Before considering the linear dependence of $n$ functions, it will be convenient to prove the following theorem for two functions:

Theorem I. *Let $y_1$ and $y_2$ be functions of the $p$ independent variables $u_1, u_2, \cdots, u_p$ for which all partial derivatives of the first order, $\partial y_1/\partial u_k$, $\partial y_2/\partial u_k$, $(k = 1, 2, \cdots, p)$ exist throughout the region $A$. Suppose, further, that one of the functions, say $y_1$, vanishes at no point of $A$. Then if all the two-rowed determinants in the matrix*

$$\left\| \begin{matrix} y_1, & y_2 \\ \dfrac{\partial y_1}{\partial u_1}, & \dfrac{\partial y_2}{\partial u_1} \\ \dfrac{\partial y_1}{\partial u_2}, & \dfrac{\partial y_2}{\partial u_2} \\ \cdot \quad \cdot & \cdot \quad \cdot \\ \cdot \quad \cdot & \cdot \quad \cdot \\ \dfrac{\partial y_1}{\partial u_p}, & \dfrac{\partial y_2}{\partial u_p} \end{matrix} \right\|$$

*vanish identically in $A$, $y_1$ and $y_2$ are linearly dependent in $A$, and in fact*

$$y_2 \equiv cy_1.$$

For, if

$$y_1 \frac{\partial y_2}{\partial u_k} - y_2 \frac{\partial y_1}{\partial u_k} \equiv 0 \qquad (k = 1, 2, \cdots, p),$$

we have on dividing by $y_1^2$,

$$\frac{\partial}{\partial u_k}\left(\frac{y_2}{y_1}\right) \equiv 0 \qquad (k = 1, 2, \cdots, p),$$

whence

$$\frac{y_2}{y_1} \equiv c, \qquad y_2 = cy_1,$$

where $c$ is a constant.

We might of course have made the statement of the hypothesis somewhat weaker, at the sacrifice of smoothness of statement, since all the two-rowed determinants of the matrix vanish if all those vanish in which the first row of the matrix constitutes a row.

We note at once that in the case of functions of more than one variable there does not exist a single determinant which may properly be called the wronskian. We might of course speak of the two-rowed determinants of the above matrix as wronskians, but in the case of more than two dependent variables, which we shall now consider, new difficulties arise. First of all, there is no connection between the number of functions and the number of

independent variables. Moreover, the various partial derivatives of the functions may exist in a very unsymmetric way, that is, a function may for instance be analytic in some of the variables, and may have only a few derivatives in the others. Consequently, there is not even a set of determinants, the structure of which we may describe once for all, which might take the place of the single wronskian in the case of functions of one variable. As a matter of fact, we shall see presently that very useful sufficient conditions for linear dependence may be stated, in which the partial derivatives with respect to some of the variables need exist only to the first order, while those with respect to other variables may be of higher orders.

We proceed to the fundamental theorem, of which Theorem I is a special case. Let $y_1, y_2, \cdots, y_n$ be functions of the $p$ independent variables $u_1, u_2, \cdots, u_p$. Let us denote by $y_i^{(1)}, y_i^{(2)}$, etc., partial derivatives of $y_i$, of any order or kind whatever. As a matter of fact, it will not be necessary to specify just what derivative of $y_i$ is denoted by the symbol $y_i^{(j)}$. If each one of the set of functions $y_i$ $(i = 1, 2, \cdots, n)$ possesses a certain partial derivative, we shall say that *the set of functions possesses that derivative.* Thus, if $\partial^2 y_i/\partial u_1 \partial u_2$, $(i = 1, 2, \cdots, n)$ exist, the set of functions possesses the derivative $\partial^2 y/\partial u_1 \partial u_2$.

We shall frequently have occasion to consider matrices of the form

$$\left\| \begin{matrix} y_1, & y_2, & \cdots, & y_r \\ y_1^{(1)}, & y_2^{(1)}, & \cdots, & y_r^{(1)} \\ y_1^{(2)}, & y_2^{(2)}, & \cdots, & y_r^{(2)} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ y_1^{(s)}, & y_2^{(s)}, & \cdots, & y_r^{(s)} \end{matrix} \right\|,$$

in which *the first row consists of the functions* $y_1, y_2, \cdots, y_r$, and the other $s$ rows of derivatives of these functions. We shall denote throughout the matrix just written by the symbol

$$M_s(y_1, y_2, \cdots, y_r).$$

We may now state the fundamental theorem:

THEOREM II. *Let the set of n functions* $y_1, y_2, \cdots, y_n$ *of the p independent variables* $u_1, u_2, \cdots, u_p$ *possess enough partial derivatives, of any orders whatever, to form a matrix*

$$M \equiv M_{n-2}(y_1, y_2, \cdots, y_n),$$

*of n columns and* $n - 1$ *rows, in which at least one of the* $(n - 1)$*-rowed determinants, say*

$$W_n \equiv \begin{vmatrix} y_1, & y_2, & \cdots, & y_{n-1} \\ y_1^{(1)}, & y_2^{(1)}, & \cdots, & y_{n-1}^{(1)} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ y_1^{(n-2)}, & y_2^{(n-2)}, & \cdots, & y_{n-1}^{(n-2)} \end{vmatrix},$$

*vanishes nowhere in $A$. Suppose, further, that all of the first derivatives of each of the elements of the above matrix $M$ exist, and adjoin to the matrix $M$ such of these derivatives as do not already appear in $M$, to form the new matrix*

$$M' \equiv M_q(y_1, y_2, \cdots, y_n),$$

*which has $n$ columns and at least $n$ rows, so that $q \geqq n-1$. Then if all the $n$-rowed determinants of the matrix $M'$ in which the determinant $W_n$ is a first minor vanish identically in $A$, the functions $y_1, y_2, \cdots, y_n$ are linearly dependent in $A$, and in fact*

$$y_n \equiv c_1 y_1 + c_2 y_2 + \cdots + c_{n-1} y_{n-1},$$

*the $c$'s being constants.*

Let us denote by $W_1, W_2, \cdots, W_n$ the $(n-1)$-rowed determinants of the matrix $M$, in such wise that $W_i$ is the determinant which does not contain the $i$th column of $M$. If we write $y_i^{(0)} = y_i$, we have by the hypothesis

$$W_1 y_1^{(j)} + W_2 y_2^{(j)} + \cdots + W_n y_n^{(j)} \equiv 0 \quad (j = 0, 1, \cdots, q). \tag{1}$$

Differentiating each of the first $n-1$ of these equations with respect to $u_k$, we obtain the equations

$$\frac{\partial W_1}{\partial u_k} y_1^{(j)} + \frac{\partial W_2}{\partial u_k} y_2^{(j)} + \cdots + \frac{\partial W_n}{\partial u_k} y_n^{(j)} + W_1 \frac{\partial y_1^{(j)}}{\partial u_k} + W_2 \frac{\partial y_2^{(j)}}{\partial u_k} + \cdots + W_n \frac{\partial y_n^{(j)}}{\partial u_k} \equiv 0 \quad (j = 0, 1, \cdots, n-2). \tag{2}$$

But since all of the first derivatives of the elements of the matrix $M$ occur in the matrix $M'$, the last part of each of the relations (2) vanishes because of (1), so that

$$\frac{\partial W_1}{\partial u_k} y_1^{(j)} + \frac{\partial W_2}{\partial u_k} y_2^{(j)} + \cdots + \frac{\partial W_n}{\partial u_k} y_n^{(j)} \equiv 0 \quad (j = 0, 1, \cdots, n-2).$$

Multiplying the first of these by the first minor of $y_1^{(0)}$ $(\equiv y_1)$ in $W_n$, the second by the first minor of $y_1^{(1)}$ in $W_n$, etc., the last by the first minor of $y_1^{(n-2)}$ in $W_n$, and adding the results, we find that

$$\frac{\partial W_1}{\partial u_k} W_n - \frac{\partial W_n}{\partial u_k} W_1 \equiv 0.$$

But this holds for every $k$; hence by Theorem I, since $W_n \neq 0$ anywhere in $A$, we have

$$W_1 = -c_1 W_n.$$

In like manner,

$$W_2 = -c_2 W_n,$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$W_{n-1} = -c_{n-1} W_n,$$

the $c$'s being constants. Consequently, from the identity

$$W_1 y_1 + W_2 y_2 + \cdots + W_n y_n \equiv 0,$$

we have

$$W_n(-c_1 y_1 - c_2 y_2 - \cdots - c_{n-1} y_{n-1} + y_n) \equiv 0.$$

Since $W_n$ is nowhere zero in $A$, it follows that

$$y_n \equiv c_1 y_1 + c_2 y_2 + \cdots + c_{n-1} y_{n-1},$$

whence our theorem.*

It should be noted that if, as stated in the hypothesis, all the $n$-rowed determinants of $M'$ of which the determinant $W_n$ is a first minor vanish, then *all the n-rowed determinants of M′ vanish.* This follows from the non-vanishing of $W_n$.†

As an example of the application of Theorem II, consider the case of four functions of two independent variables, say

$$y_i = f_i(u, v) \qquad (i = 1, 2, 3, 4).$$

Then if in the matrix

$$\left\| \begin{matrix} y_1, & y_2, & y_3, & y_4 \\ \dfrac{\partial y_1}{\partial u}, & \dfrac{\partial y_2}{\partial u}, & \dfrac{\partial y_3}{\partial u}, & \dfrac{\partial y_4}{\partial u} \\ \dfrac{\partial y_1}{\partial v}, & \dfrac{\partial y_2}{\partial v}, & \dfrac{\partial y_3}{\partial v}, & \dfrac{\partial y_4}{\partial v} \end{matrix} \right\|$$

at least one of the three-rowed determinants vanishes nowhere in the region $A$, a necessary and sufficient condition for the linear dependence of $y_1$, $y_2$, $y_3$, $y_4$ is the identical vanishing in $A$ of all the four-rowed determinants of

* This proof is a direct generalization of the corresponding one given by Frobenius for functions of a single variable, in the case of the wronskian. Cf. Bôcher, *Certain cases in which the vanishing of the wronskian is a sufficient condition for linear dependence*, these Transactions, vol. 2 (1901), pp. 139–149, § 1.

† Cf. Bôcher, *Introduction to Higher Algebra*, p. 54, for the general theorem on matrices involved here.

the matrix formed by adjoining to the above one all of the derivatives of the second order of $y_1, y_2, y_3, y_4$. If, however, at least one of the three-rowed determinants in the matrix

$$\left\| \begin{matrix} y_1, & y_2, & y_3, & y_4 \\ \dfrac{\partial y_1}{\partial u}, & \dfrac{\partial y_2}{\partial u}, & \dfrac{\partial y_3}{\partial u}, & \dfrac{\partial y_4}{\partial u} \\ \dfrac{\partial^2 y_1}{\partial u^2}, & \dfrac{\partial^2 y_2}{\partial u^2}, & \dfrac{\partial^2 y_3}{\partial u^2}, & \dfrac{\partial^2 y_4}{\partial u^2} \end{matrix} \right\|$$

vanishes nowhere in $A$, then a necessary and sufficient condition for the linear dependence of $y_1, y_2, y_3, y_4$ is the identical vanishing of all the four-rowed determinants of the matrix

$$\left\| \begin{matrix} y_1, & \dfrac{\partial y_1}{\partial u}, & \dfrac{\partial y_1}{\partial v}, & \dfrac{\partial^2 y_1}{\partial u^2}, & \dfrac{\partial^2 y_1}{\partial u \partial v}, & \dfrac{\partial^3 y_1}{\partial u^3}, & \dfrac{\partial^3 y_1}{\partial u^2 \partial v} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ y_4, & \dfrac{\partial y_4}{\partial u}, & \dfrac{\partial y_4}{\partial v}, & \dfrac{\partial^2 y_4}{\partial u^2}, & \dfrac{\partial^2 y_4}{\partial u \partial v}, & \dfrac{\partial^3 y_4}{\partial u^3}, & \dfrac{\partial^3 y_4}{\partial u^2 \partial v} \end{matrix} \right\|,$$

so that the existence of even $\partial^2 y_i/\partial v^2$ is not presupposed.

In the fundamental theorem concerning the linear dependence of $n$ functions of a single variable, the identical vanishing of the wronskian is shown to be sufficient for linear dependence provided the wronskian of $n-1$ of the functions is not zero anywhere in the interval of the independent variable under consideration. This evidently comes as a special case under Theorem II. But Theorem II also affords other sufficient criteria in the case of a single independent variable. An example will show the general nature of these criteria. Let $y, z, w$ be functions of the real variable $x$ in an interval $I$. Denoting by accents differentiation with respect to $x$, let us suppose that the determinant

$$\begin{vmatrix} y, & z \\ y''', & z''' \end{vmatrix}$$

vanishes nowhere in $I$. Then a sufficient condition for the linear dependence of $y, z, w$ is the identical vanishing in $I$ of the two determinants

$$\begin{vmatrix} y, & z, & w \\ y', & z', & w' \\ y''', & z''', & w''' \end{vmatrix}, \qquad \begin{vmatrix} y, & z, & w \\ y''', & z''', & w''' \\ y^{\text{IV}}, & z^{\text{IV}}, & w^{\text{IV}} \end{vmatrix}.$$

However, in any case such as the present one, a very general theorem concerning linear dependence, due to D. R. Curtiss, is applicable.*

Before leaving the fundamental theorem, it may be well to consider the case of analytic functions. For $n$ analytic functions of a single variable, the identical vanishing of the wronskian is a sufficient condition for linear dependence, and nothing need be demanded as to the non-vanishing of the wronskian of $n-1$ of the functions. This follows from the fact that for analytic functions linear dependence in any subregion implies linear dependence throughout their common region of definition. If in Theorem II the functions involved are analytic, an obvious weakening of the hypotheses may be made; for instance, it is sufficient to know that *all the n-rowed determinants of the matrix $M'$ which contain the elements of $W_n$ vanish throughout an n-dimensional sub-region of $A$ which contains a single point at which $W_n$ is different from zero.* Such would be the most general theorem that may be stated in this case; however, by strengthening the hypotheses sufficiently we may obtain a theorem in which no demand is made concerning the non-vanishing of $W_n$. A very elegant theorem of this kind is the following, which is due to O. D. Kellogg:†

*Let $y_1, y_2, \cdots, y_n$ be analytic functions of the p independent variables $u_1, u_2, \cdots, u_p$—real or complex—having a common region of definition $A$, and let $M_q(y_1, y_2, \cdots, y_n)$ be the matrix formed from the functions $y_1, y_2, \cdots, y_n$ and all their partial derivatives up to and including those of order $n-1$. Then the functions $y_1, y_2, \cdots, y_n$ are linearly dependent in $A$ if and only if all of the n-rowed determinants of the matrix $M_q$ vanish throughout $A$.*

## 2. Two theorems concerning the linear dependence of $n$ functions

In the present section we shall prove two theorems, which are generalizations of theorems given by Bôcher‡ for functions of a single variable. They will be useful later, in connection with completely integrable systems of partial differential equations, just as Bôcher's theorems are in the case of ordinary linear differential equations, of which completely integrable systems are a generalization to several independent variables.

Theorem III. *Let $y_1, y_2, \cdots, y_n$ be n functions of the variables $u_1, u_2, \cdots, u_p$. Suppose that the set of functions possesses the $n-1$ partial derivatives $y_i^{(1)}, y_i^{(2)}, \cdots, y_i^{(n-1)}$, $(i = 1, 2, \cdots, n)$, so that the function*

* Cf. Theorem VI, p. 292, of his article, *The vanishing of the wronskian and the problem of linear dependence*, Mathematische Annalen, vol. 65 (1908), pp. 282–298.

† *Nomograms with points in alignment*, Zeitschrift für Mathematik und Physik, vol. 63 (1914), pp. 160–161. The statement there given is equivalent to ours, if the proper interpretation is made of Prof. Kellogg's definition of the rank of a matrix.

‡ Cf. Theorems V and VI of the paper already cited, these Transactions, vol. 2 (1901), pp. 139–149.

$$y = c_1 y_1 + c_2 y_2 + \cdots + c_n y_n,$$

*where the c's are constants, has the same derivatives,* $y^{(1)}, y^{(2)}, \cdots, y^{(n-1)}$. *Suppose also that no function of the form of* $y$, *other than zero, vanishes together with its derivatives* $y^{(1)}, y^{(2)}, \cdots, y^{(n-1)}$ *at any point of* $A$. *Then if the determinant*

$$W \equiv \begin{vmatrix} y_1, & y_2, & \cdots, & y_n \\ y_1^{(1)}, & y_2^{(1)}, & \cdots, & y_n^{(1)} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ y_1^{(n-1)}, & y_2^{(n-1)}, & \cdots, & y_n^{(n-1)} \end{vmatrix}$$

*vanishes at any point* $(a_1, a_2, \cdots, a_p)$ *of* $A$, *the functions* $y_1, y_2, \cdots, y_n$ *are linearly dependent.*

For, since $W(a_1, a_2, \cdots, a_p) = 0$, there exist $n$ constants $c_1, c_2, \cdots, c_n$ not all zero, such that

$$c_1 y_1^{(j)}(a_1, \cdots, a_p) + c_2 y_2^{(j)}(a_1, \cdots, a_p) + \cdots + c_n y_n^{(j)}(a_1, \cdots, a_p) = 0 \qquad (j = 0, 1, \cdots, n-1).$$

so that the function $y = c_1 y_1 + c_2 y_2 + \cdots + c_n y_n$ vanishes together with its derivatives $y^{(1)}, y^{(2)}, \cdots, y^{(n-1)}$ at a point of $A$, and must therefore be identically zero.

THEOREM IV. *Let* $y_1, y_2, \cdots, y_n$ *be a set of* $n$ *functions of the* $p$ *variables* $u_1, u_2, \cdots, u_p$ *which throughout* $A$ *possesses* $r$ *partial derivatives* $(r > n-1)$,

$$y_i^{(1)}, y_i^{(2)}, \cdots, y_i^{(r)} \qquad (i = 1, 2, \cdots, n).$$

*Suppose that from among these derivatives may be selected at least* $n-2$, *say* $y_i^{(1)}, y_i^{(2)}, \cdots, y_i^{(m)}$ $(m \geqq n-2)$, *such that all of their first derivatives occur among the set* $y_i^{(1)}, y_i^{(2)}, \cdots, y_i^{(r)}$, *and are bounded in* $A$. *Let*

$$M \equiv M_m(y_1, y_2, \cdots, y_n) \qquad (m \geqq n-2),$$

*and suppose that*

$$M' \equiv M_q(y_1, y_2, \cdots, y_n) \qquad (q > m)$$

*is the matrix formed from* $M$ *by adjoining to the elements of* $M$ *those of their first derivatives which do not already appear in* $M$. *Assume also that if for any* $\nu$ *columns of* $M$ $(\nu \leqq n)$ *all the* $\nu$*-rowed determinants containing elements of the first row of* $M$ *vanish identically, then all the* $\nu$*-rowed determinants formed from the same* $\nu$ *columns of* $M'$, *and containing elements of the first row of* $M$, *and of* $\nu - 2$ *other rows of* $M$, *vanish identically.** *Suppose, further, that no function, other than zero, of the form*

* Some remarks concerning this part of the hypothesis will be found after the proof of the theorem.

$$y = c_1 y_1 + c_2 y_2 + \cdots + c_n y_n,$$

*where the c's are constants, vanishes together with its derivatives* $y^{(1)}, y^{(2)}, \cdots, y^{(r)}$ *at any point of* $A$. *Then the identical vanishing of all those n-rowed determinants of* $M'$ *in which appear elements of the first row and* $n - 2$ *other rows of* $M$ *is a sufficient condition that the functions* $y_1, y_2, \cdots, y_n$ *be linearly dependent.*

Suppose first that one of the $(n-1)$-rowed determinants of the matrix $M$ containing elements of the first row of $M$ does not vanish identically. We may without loss of generality take for this determinant the one which appears in the upper left-hand corner of $M$. As in § 1, we denote this determinant by $W_n$. Then at some point $(a_1, a_2, \cdots, a_p)$ of $A$, $W_n$ is different from zero. But $W_n$ is continuous in $A$, because all its partial derivatives of the first order exist and are bounded in $A$*; consequently, it differs from zero in some neighborhood of the point $(a_1, a_2, \cdots, a_p)$. By Theorem II, then, the functions $y_1, y_2, \cdots, y_n$ are linearly dependent in this neighborhood, i. e., there exist constants $c_1, c_2, \cdots, c_n$, not all zero, such that the function

$$y = c_1 y_1 + c_2 y_2 + \cdots + c_n y_n$$

is zero in the said neighborhood of $(a_1, a_2, \cdots, a_p)$. But then this function vanishes at $(a_1, a_2, \cdots, a_p)$ together with its derivatives $y^{(1)}, y^{(2)}, \cdots, y^{(r)}$, and therefore, by hypothesis, vanishes identically. Consequently the functions $y_1, y_2, \cdots, y_n$ are linearly dependent in $A$.

We now assume the other alternative, viz., that all of the $(n-1)$-rowed determinants of $M$ which contain elements of the first row of $M$ vanish identically; in particular, those formed from the first $n-1$ columns of $M$ do. We recall that, by hypothesis, if all the $\nu$-rowed determinants formed from the first $\nu$ columns of $M$ and containing elements of the first row of $M$ vanish identically, then the determinants formed from the first $\nu$ columns of $M'$, and containing elements of the first row and of $\nu - 2$ other rows of $M$, vanish identically. We consider then in succession the first $n-1$ columns of $M$, the first $n-2$ columns, the first $n-3$ columns, etc. Suppose that for the first $n-i$ columns of $M$ all of the $(n-i)$-rowed determinants which contain elements of the first row of $M$ vanish identically. We know that this happens for $i = 0$ and $i = 1$. If it happens for $i = n-1$, we shall have $y_1 \equiv 0$, and then $y_1, y_2, \cdots, y_n$ would be linearly dependent. Suppose then that $i < n-1$. Then there must exist a particular $i$, $1 \leqq i \leqq n-2$, such that all the $(n-i)$-rowed determinants formed from the first $n-i$ columns of $M$ and containing elements of the first row of $M$ vanish identically, whereas

* If for a function of $p$ variables all the partial derivatives of the first order exist and are bounded in the $p$-dimensional region $A$, the function is continuous in $A$ with respect to all the variables taken together. Cf. Osgood, *Lehrbuch der Funktionentheorie*, vol. I, second edition (1912), page 56.

the $(n-i-1)$-rowed determinants formed from the first $n-i-1$ columns of $M$ and containing elements of the first row of $M$ do not all vanish identically. But the identical vanishing of all the $(n-i)$-rowed determinants of the kind considered, formed from the first $n-i$ columns of $M$, carries with it, by hypothesis, the identical vanishing of all the $(n-i)$-rowed determinants formed from the first $n-i$ columns of $M'$ and containing elements of the first row and of $n-i-2$ other rows of $M$. We may therefore apply the special case of the present theorem which we have already proved, to deduce the linear dependence of $y_1, y_2, \cdots, y_{n-i}$. Consequently, the functions $y_1, y_2, \cdots, y_n$ are also linearly dependent, and our theorem is proved.

It is to be noted that in the hypothesis of Theorem IV, a restriction has been placed upon the nature of the matrices $M$ and $M'$ which we did not find it necessary to impose in the previous general theorems. As may readily be seen by an example, it is only in special cases that the identical vanishing of all the $\nu$-rowed determinants formed from $\nu$ columns of the matrix $M$ and containing elements of the first row of $M$ implies the identical vanishing of all the $\nu$-rowed determinants formed from the corresponding $\nu$ columns of $M'$ and containing elements of the first row and of $\nu-2$ additional rows of $M$. Probably the most important cases of this type arise when the functions $y_1, y_2, \cdots, y_n$ are solutions of the same completely integrable system of partial differential equations. We shall have occasion, in a subsequent section, to apply Theorem IV in just such a case.

## 3. Completely integrable systems of homogeneous linear partial differential equations

A system of partial differential equations in one dependent variable is said to be *completely integrable* if there exists one and only one integral of the system satisfying arbitrarily chosen initial conditions involving a finite number of arbitrary constants and arbitrary functions. The completely integrable systems which we shall consider are such that a general solution is expressible linearly, with constant coefficients, in terms of a fundamental system of solutions. Thus, these completely integrable systems, which we shall suppose to contain a single dependent variable and any number of independent variables, form a natural generalization of the ordinary linear homogeneous differential equation.

So far as we are aware, the only completely integrable systems containing partial differential equations of order higher than the first which have hitherto been studied have analytic coefficients and solutions.* From the function-theoretic point of view these are of course the most important completely

* Cf., for example, C. Riquier, *Les systèmes d'équations aux dérivées partielles*, Paris, Gauthier-Villars, 1910.

integrable systems. They have found a place not only in analysis, but in geometry as well. The Gauss-Codazzi equations which form the basis for the metric differential geometry of surfaces constitute a completely integrable system, and in a later section we shall see how important are such systems in projective differential geometry.

The simplicity of the analytic case, if we disregard questions of convergence, is apparent. Let us consider, as always, a completely integrable system of linear homogeneous partial differential equations containing a single dependent variable; then by successive differentiation of the given equations, all the derivatives of the dependent variable are expressible uniquely as linear combinations of a finite number $n$ among them, so that by well-known methods it is seen that the general solution contains exactly the same number $n$ of arbitrary constants, which enter linearly.

We shall confine our attention throughout to linear homogeneous completely integrable systems in one dependent variable, though most of what we shall have to say may readily be extended to the case of several dependent variables. We shall not, however, suppose the coefficients in our equations to be analytic.*

Let us first construct a completely integrable system which shall have as a fundamental system of solutions an assigned set of functions. Suppose that the $n$ functions $y_1, y_2, \cdots, y_n$ of the $p$ independent variables $u_1, u_2, \cdots, u_p$ possess enough partial derivatives, of any kind or order whatever, to form a determinant of the $n$th order,

$$(3) \qquad W \equiv \begin{vmatrix} y_1, & y_2, & \cdots, & y_n \\ y_1^{(1)}, & y_2^{(1)}, & \cdots, & y_n^{(1)} \\ y_1^{(2)}, & y_2^{(2)}, & \cdots, & y_n^{(2)} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ y_1^{(n-1)}, & y_2^{(n-1)}, & \cdots, & y_n^{(n-1)} \end{vmatrix},$$

which vanishes nowhere in a region $A$, and in which, as heretofore, we have denoted derivatives of a function $y$ by $y^{(1)}$, $y^{(2)}$, etc., the indices being no indication of the kind or order of the derivative. Suppose also that all of the derivatives of the first order exist for all of the elements in the determinant

* Completely integrable systems in which the coefficients are not restricted to being analytic were curiously enough the first to be studied, but only such systems as consist of equations of the first order. These systems are really systems of total differential equations, and the classic work of Mayer in connection therewith is well known. See his memoir, *Ueber unbeschränkt integrable Systeme von linearen totalen Differentialgleichungen und die simultane Integration linearer partieller Differentialgleichungen*, Mathematische Annalen, vol. 5 (1872), pp. 448–470.

$W$. Then we may set up as follows a system of partial differential equations, of which each of the functions $y_1, y_2, \cdots, y_n$ is a solution, and such that any solution whatever, say $\bar{y}$, is of the form

$$\bar{y} \equiv c_1 y_1 + c_2 y_2 + \cdots + c_n y_n,$$

where the $c$'s are constants. Let us write down the system of $np$ equations

$$\begin{aligned}
\frac{\partial y}{\partial u_k} &= a_0^{(0,k)} y + a_1^{(0,k)} y^{(1)} + a_2^{(0,k)} y^{(2)} + \cdots + a_{n-1}^{(0,k)} y^{(n-1)}, \\
\frac{\partial y^{(1)}}{\partial u_k} &= a_0^{(1,k)} y + a_1^{(1,k)} y^{(1)} + a_2^{(1,k)} y^{(2)} + \cdots + a_{n-1}^{(1,k)} y^{(n-1)}, \\
\frac{\partial y^{(2)}}{\partial u_k} &= a_0^{(2,k)} y + a_1^{(2,k)} y^{(1)} + a_2^{(2,k)} y^{(2)} + \cdots + a_{n-1}^{(2,k)} y^{(n-1)}, \qquad (4) \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
\frac{\partial y^{(n-1)}}{\partial u_k} &= a_0^{(n-1,k)} y + a_1^{(n-1,k)} y^{(1)} + a_2^{(n-1,k)} y^{(2)} + \cdots + a_{n-1}^{(n-1,k)} y^{(n-1)}, \\
&\qquad (k = 1, 2, \cdots, p).
\end{aligned}$$

We may so determine the coefficients $a_i^{(j,k)}$ of this system that the functions $y_1, y_2, \cdots, y_n$ will be solutions of the system. If in any one of the equations these functions be substituted successively, $n$ linear equations are obtained which may be solved for the coefficients $a$ of the first mentioned equation since the determinant $W$ is different from zero. If the coefficients $a$ are determined in this way, the system of differential equations will surely have as solutions each of the functions $y_1, y_2, \cdots, y_n$. In fact, the system of $np$ differential equations may be written in the compact form

$$(5) \qquad \begin{vmatrix} y_1, & y_2, & \cdots, & y_n, & y \\ y_1^{(1)}, & y_2^{(1)}, & \cdots, & y_n^{(1)}, & y^{(1)} \\ y_1^{(2)}, & y_2^{(2)}, & \cdots, & y_n^{(2)}, & y^{(2)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ y_1^{(n-1)}, & y_2^{(n-1)}, & \cdots, & y_n^{(n-1)}, & y^{(n-1)} \\ \dfrac{\partial y_1^{(j)}}{\partial u_k}, & \dfrac{\partial y_2^{(j)}}{\partial u_k}, & \cdots, & \dfrac{\partial y_n^{(j)}}{\partial u_k}, & \dfrac{\partial y^{(j)}}{\partial u_k} \end{vmatrix} = 0$$

$$(j = 0, 1, \cdots, n-1;\ k = 1, 2, \cdots, p),$$

which makes the statement evident. Moreover, it is also evident that any expression of the form

$$\bar{y} = c_1 y_1 + c_2 y_2 + \cdots + c_n y_n$$

will also satisfy the system. Lastly, any solution of the system must be of this form. For, let $\bar{y}$ be any solution of the system, and consider the two matrices

$$M \equiv M_{n-1}(y_1, y_2, \cdots, y_n, y),$$
$$M' \equiv M_q(y_1, y_2, \cdots, y_n, y), \tag{6}$$

where $M'$ differs from $M$ by the additional rows consisting of all of the partial derivatives of the first order of the elements of $M$ which do not already appear in $M$. The derivatives appearing in these additional rows are precisely those which appear in the left-hand members of the system of differential equations (4), so that each of the additional rows is a linear combination of the $n$ rows of the matrix $M$. Consequently, all of the $(n+1)$-rowed determinants of $M'$ which contain the elements of $M$ vanish identically. But one of the $n$-rowed determinants of $M$, namely $W$, vanishes nowhere in $A$, so that by Theorem II the $n+1$ functions $y_1, y_2, \cdots, y_n, \bar{y}$ are linearly dependent, and in fact

$$\bar{y} = c_1 y_1 + c_2 y_2 + \cdots + c_n y_n.$$

We may state our result as follows:

THEOREM V. *If the system of partial differential equations* (4) *has n solutions* $y_1, y_2, \cdots, y_n$ *for which the determinant* (3) *vanishes nowhere in* $A$, *then any solution of the system will be linearly dependent on these n solutions, and any function linearly dependent on these n solutions will also be a solution.*

We have therefore obtained a completely integrable system of which the given functions $y_1, y_2, \cdots, y_n$ constitute a fundamental system of solutions. The form in which we have chosen to write the system of differential equations is not generally the simplest for any given particular case; in fact, in many cases most of the equations will be found to be merely identities, and others of them obtainable from the rest by mere differentiation. The completely integrable system will in such a case reduce to fewer than $np$ equations. An example will make the point clearer. Suppose that the four functions

$$y_k = y_k(u_1, u_2) \qquad (k = 1, 2, 3, 4)$$

are such that the determinant

$$W = \begin{vmatrix} y_1, & y_2, & y_3, & y_4 \\ \dfrac{\partial y_1}{\partial u_1}, & \dfrac{\partial y_2}{\partial u_1}, & \dfrac{\partial y_3}{\partial u_1}, & \dfrac{\partial y_4}{\partial u_1} \\ \dfrac{\partial y_1}{\partial u_2}, & \dfrac{\partial y_2}{\partial u_2}, & \dfrac{\partial y_3}{\partial u_2}, & \dfrac{\partial y_4}{\partial u_2} \\ \dfrac{\partial^2 y_1}{\partial u_1 \partial u_2}, & \dfrac{\partial^2 y_2}{\partial u_1 \partial u_2}, & \dfrac{\partial^2 y_3}{\partial u_1 \partial u_2}, & \dfrac{\partial^2 y_4}{\partial u_1 \partial u_2} \end{vmatrix}$$

is zero at no point of the region $A$. Then the completely integrable system of which the four functions $y_1, y_2, y_3, y_4$ form a fundamental system of solutions will be

$$\frac{\partial y}{\partial u_1} = \frac{\partial y}{\partial u_1}, \qquad \frac{\partial y}{\partial u_2} = \frac{\partial y}{\partial u_2}, \qquad \frac{\partial^2 y}{\partial u_1 \partial u_2} = \frac{\partial^2 y}{\partial u_1 \partial u_2},$$

$$\frac{\partial^2 y}{\partial u_1^2} = ay + b\frac{\partial y}{\partial u_1} + c\frac{\partial y}{\partial u_2} + d\frac{\partial^2 y}{\partial u_1 \partial u_2},$$

$$\frac{\partial^2 y}{\partial u_2^2} = a' y + b'\frac{\partial y}{\partial u_1} + c'\frac{\partial y}{\partial u_2} + d'\frac{\partial^2 y}{\partial u_1 \partial u_2},$$

$$\frac{\partial^3 y}{\partial u_1^2 \partial u_2} = a'' y + b''\frac{\partial y}{\partial u_1} + c''\frac{\partial y}{\partial u_2} + d''\frac{\partial^2 y}{\partial u_1 \partial u_2},$$

$$\frac{\partial^3 y}{\partial u_1 \partial u_2^2} = a''' y + b'''\frac{\partial y}{\partial u_1} + c'''\frac{\partial y}{\partial u_2} + d'''\frac{\partial^2 y}{\partial u_1 \partial u_2};$$

we have of course presupposed the existence of all the derivatives which we have written down. The first three of the above equations are identities, and as a matter of fact the third of them ought to occur twice. The completely integrable system therefore consists essentially of four instead of $2 \times 4 = 8$ equations. Moreover, if $dd' - 1$ is nowhere zero in $A$, the last two of the differential equations may be obtained from the two immediately preceding them by differentiation and algebraic processes. The completely integrable system of eight equations in this case reduces therefore to the two equations

$$\frac{\partial^2 y}{\partial u_1^2} = ay + b\frac{\partial y}{\partial u_1} + c\frac{\partial y}{\partial u_2} + d\frac{\partial^2 y}{\partial u_1 \partial u_2},$$

$$\frac{\partial^2 y}{\partial u_2^2} = a' y + b'\frac{\partial y}{\partial u_1} + c'\frac{\partial y}{\partial u_2} + d'\frac{\partial^2 y}{\partial u_1 \partial u_2}.$$

It is evident that in any particular problem this form of the system is most useful; however, in the very general questions which we shall treat subsequently, the form (4) in which we first wrote the system seems of great assistance, in spite of the fact that many of the equations may be superfluous.

From what we have already seen, the analogy of homogeneous linear completely integrable systems to ordinary homogeneous linear differential equations in very close. In later sections the similarity will be made to appear even more striking; at present we take occasion to make the following remark. If in system (4) we keep $k$ fixed, we have a system of $n$ differential equations which may be regarded as a system of ordinary differential equations in which the independent variable is $u_k$ and the dependent variables are $n$ in number, viz., $y, y^{(1)}, y^{(2)}, \cdots, y^{(n-1)}$. *The completely integrable system* (4) *may there-*

*fore be regarded in $p$ ways as a system of $n$ ordinary differential equations in $n$ dependent variables, containing $p-1$ parameters.*

## 4. The integrability conditions and wronskian of a completely integrable system

In the last section, the completely integrable system (4) was set up on the supposition that a fundamental system of solutions was known. Naturally, if a system of this form, which we shall write

$$\frac{\partial y^{(j)}}{\partial u_k} = \sum_{i=0}^{n-1} a_i^{(j,k)} y^{(i)} \quad (j = 0, 1, \cdots, n-1;\ k = 1, 2, \cdots, p), \tag{7}$$

is to be completely integrable and have precisely $n$ fundamental solutions, the coefficients cannot be arbitrary. We shall in the present section investigate certain relations which exist among the coefficients of system (7), under the supposition that the system has a fundamental system of $n$ solutions $y_1, y_2, \cdots, y_n$ in the region $A$.

Let us call the quantities $y, y^{(1)}, y^{(2)}, \cdots, y^{(n-1)}$ which occur in the right-hand members of equations (7), the *primary derivatives* of the function $y$.

The determinant $W$ given by equation (3) is formed of the primary derivatives of a fundamental system of solutions of the differential equations. We shall call this determinant the *wronskian* of the system of solutions. In setting up the completely integrable system (4) we assumed that the wronskian of the fundamental system of solutions vanished nowhere in $A$; this is equivalent to the assumption that *if for every solution of the system of differential equations a relation exists among the primary derivatives, of the form*

$$\alpha_0 y + \alpha_1 y^{(1)} + \alpha_2 y^{(2)} + \cdots + \alpha_{n-1} y^{(n-1)} \equiv 0, \tag{8}$$

*then all of the functions $\alpha_0, \alpha_1, \cdots, \alpha_{n-1}$ must be identically zero.*

We shall from now on assume that in the region $A$ the coefficients $a_i^{(j,k)}$ of system (7) are bounded and possess all partial derivatives of the first order. Then since by supposition there exists a function $y$ not identically zero which satisfies equations (7), we may differentiate these equations and obtain in two ways an expression for $\partial^2 y^{(j)}/\partial u_k \partial u$, $(k \neq l)$. In fact,

$$\frac{\partial^2 y^{(j)}}{\partial u_k \partial u_l} = \sum_{i=0}^{n-1} \left( \frac{\partial a_i^{(j,k)}}{\partial u_l} y^{(i)} + a_i^{(j,k)} \frac{\partial y^{(i)}}{\partial u_l} \right),$$

$$\frac{\partial^2 y^{(j)}}{\partial u_l \partial u_k} = \sum_{i=0}^{n-1} \left( \frac{\partial a_i^{(j,l)}}{\partial u_k} y^{(i)} + a_i^{(j,l)} \frac{\partial y^{(i)}}{\partial u_k} \right).$$

In the right-hand members of each of these, the derivatives of $y^{(i)}$ may be replaced by their expressions in terms of $y, y^{(1)}, \cdots, y^{(n-1)}$ taken from sys-

tem (7). In this way two different expressions of the form (8) are obtained for a derivative $\partial^2 y^{(j)}/\partial u_k \partial u_l$ $(k \neq l)$. But these two expressions are to be identically equal, so that by our assumption the coefficients of $y^{(i)}$ in the two expressions must be identically equal. We may therefore state the theorem:

THEOREM VI. *A necessary condition that system* (7) *have a fundamental system of n solutions is that the coefficients satisfy identically the relations*

$$\frac{\partial a_\nu^{(j,k)}}{\partial u_l} + \sum_{i=0}^{n-1} a_i^{(j,k)} a_\nu^{(i,l)} = \frac{\partial a_\nu^{(j,l)}}{\partial u_k} + \sum_{i=0}^{n-1} a_i^{(j,l)} a_\nu^{(i,k)} \tag{9}$$

$$(\nu, j = 0, 1, \cdots, n-1;\ k, l = 1, 2, \cdots, p).$$

We shall speak of equations (9) as the *conditions of complete integrability*, or as the *integrability conditions*, of system (7).

From the set of integrability conditions (9) let us pick out $n$ equations, by keeping $k$ and $l$ fixed and choosing $\nu = j = 0, 1, \cdots, n-1$. The equations are

$$\frac{\partial a_j^{(j,k)}}{\partial u_l} + \sum_{i=0}^{n-1} a_i^{(j,k)} a_j^{(i,l)} = \frac{\partial a_j^{(j,l)}}{\partial u_k} + \sum_{i=0}^{n-1} a_i^{(j,l)} a_j^{(i,k)}. \quad (j = 0, 1, \cdots, n-1).$$

Adding them, we find without difficulty that all terms drop out except the first derivatives, and in fact

$$\sum_{j=0}^{n-1} \frac{\partial a_j^{(j,k)}}{\partial u_l} = \sum_{j=0}^{n-1} \frac{\partial a_j^{(j,l)}}{\partial u_k} \qquad (k, l = 1, 2, \cdots, p),$$

that is,

$$\frac{\partial}{\partial u_l} \sum_{j=0}^{n-1} a_j^{(j,k)} = \frac{\partial}{\partial u_k} \sum_{j=0}^{n-1} a_j^{(j,l)}. \tag{10}$$

Consequently, the $n$ quantities $\sum_{j=0}^{n-1} a_j^{(j,k)}$ are the first derivatives of some function $f(u_1, u_2, \cdots, u_p)$, and we may write

$$\frac{\partial f}{\partial u_k} = \sum_{j=0}^{n-1} a_j^{(j,k)} \qquad (k = 1, 2, \cdots, p). \tag{11}$$

We shall presently find relations (10) and (11) very useful.

Let us now form the wronskian (3) of any fundamental system of $n$ solutions $y_1, y_2, \cdots, y_n$ of system (7):

$$W = \begin{vmatrix} y_1, & y_2, & \cdots, & y_n \\ y_1^{(1)}, & y_2^{(1)}, & \cdots, & y_n^{(1)} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ y_1^{(n-1)}, & y_2^{(n-1)}, & \cdots, & y_n^{(n-1)} \end{vmatrix}.$$

We differentiate $W$ with respect to $u_k$; if the result be written as the sum of the $n$ determinants formed from $W$ by replacing a row thereof by the deriva-

tives of that row with respect to $u_k$, and if use be then made of equations (7) to replace any derivative $\partial y_i^{(j)}/\partial u_k$ by an expression linear in $y_i$, $y_i^{(1)}$, $\cdots$, $y_i^{(n-1)}$, it will be found without difficulty that

$$\frac{\partial W}{\partial u_k} W = \sum_{j=0}^{n-1} a_j^{(j,k)} = \frac{\partial f}{\partial u_k} W \qquad (k = 1, 2, \cdots, p).$$

Consequently,

$$W = const.\ e^f.$$

We may therefore state the theorem, which is a direct generalization of the familiar theorem of Abel in the theory of an ordinary homogeneous linear differential equation of the $n$th order:

THEOREM VII. *The wronskian $W$ of a fundamental system of $n$ solutions of the completely integrable system* (7) *may be determined by a quadrature from the coefficients of the system, and is given by the expression*

$$W = const.\ e^f, \tag{12}$$

*where $f$ may be found by a quadrature in virtue of equations* (10) *or* (11).

As a corollary we have the theorem:

THEOREM VIII. *If the wronskian vanish at any point of the region $A$, it vanishes identically in $A$.**

We remark in passing that equations (10) and (11) might have been deduced directly from a consideration of the wronskian. For, equations (10) may be deduced by differentiation of the relations

$$\frac{\partial W}{\partial u_k} = W \sum_{j=0}^{n-1} a_j^{(j,k)} \qquad (k = 1, 2, \cdots, p). \tag{13}$$

But the relations (13) are of interest in themselves, because they show that the function $\sum_{j=0}^{n-1} a_j^{(j,k)}$ is the derivative, with respect to $u_k$, of a function $f$, for $k = 1, 2, \cdots, p$, and this *without necessitating the assumption of the existence of any of the first derivatives of the coefficients $a_i^{(j,k)}$ of system* (7).

## 5. THE NORMAL FORM OF THE COMPLETELY INTEGRABLE SYSTEM

A necessary condition that a system of partial differential equations of the form (7) have a solution not identically zero is that the coefficients of the system satisfy the integrability conditions (9). Whether the condition is also sufficient remains to be seen; as a matter of fact, however, an example will show that a system of form (7), whose coefficients satisfy the integrability conditions, need not have a solution which is not identically zero. For, if we select as primary derivatives $y$ and $\partial^2 y/\partial u^2$, and suppose the independent variables to be $u$ and $v$, we may write down the system of form (7):

* We denote by $A$ the region in which all of the solutions $y_1, y_2, \cdots, y_n$ exist.

$$\frac{\partial y}{\partial u} = ay + b\frac{\partial^2 y}{\partial u^2}, \qquad \frac{\partial y}{\partial v} = a'y + b'\frac{\partial^2 y}{\partial u^2},$$

$$\frac{\partial^3 y}{\partial u^3} = a''y + b''\frac{\partial^2 y}{\partial u^2}, \qquad \frac{\partial^3 y}{\partial u^2 \partial v} = a'''y + b'''\frac{\partial^2 y}{\partial u^2},$$

which has no solution different from zero even if the integrability conditions of form (9) are satisfied for the system, unless of course other relations exist among the coefficients. In fact, differentiation of the first equation of the system twice with respect to $u$ will lead to two new equations which are in general incompatible with the first and third equations of the system. In our example this results from the peculiar choice of primary derivatives, a gap occurring between $y$ and $\partial^2 y/\partial u^2$.

We shall presently show that it is always possible, for a system of form (7) which is known to have $n$ fundamental solutions, to make a particular choice of primary derivatives which removes the above difficulty, although at a slight sacrifice in generality. Let us first define what we shall call a *chain* of derivatives and a *normal set* of derivatives. A *chain* of derivatives of a function $y$ is a set of derivatives $y, y^{(1)}, y^{(2)}, \cdots, y^{(r)}$, such that $y^{(i)}$ is of the $i$th order and may be obtained by a differentiation from $y^{(i-1)}$. Thus, denoting partial differentiation by subscripts, we see that the quantities $y, y_u, y_{uv}, y_{uuv}$ form a chain, while $y, y_u, y_{uv}, y_{uuu}$ do not because $y_{uuu}$ cannot be formed from $y_{uv}$ by differentiation.

A *normal set* of derivatives is a set such that each element of the set belongs to at least one chain which can be formed from elements of the set. The normal set need not consist of a single chain, and a given member of the set may belong to more than one chain.

We shall now show that if system (7) has exactly $n$ fundamental solutions $y_1, y_2, \cdots, y_n$, the primary derivatives may under certain conditions be so chosen as to form a normal set. However, in order not to interrupt our reasoning later, we shall first prove a lemma which is to be used in the course of the discussion.

LEMMA. *Suppose that system* (7) *has n solutions* $y_1, y_2, \cdots, y_n$ *whose wronskian W vanishes nowhere in the region A. Then* $y_1, y_2, \cdots, y_n$ *cannot all be solutions of a system R of form* (7) *having less than n primary derivatives, these primary derivatives being selected from among those of system* (7).

Let $y, y^{(1)}, y^{(2)}, \cdots, y^{(r)}$ $(r < n-1)$ be the primary derivatives of the system $R$. Then the lemma follows at once if in the matrix

$$M \equiv M_r(y_1, y_2, \cdots, y_n)$$

at least one of the $(r+1)$-rowed determinants vanishes at no point of $A$, for then this determinant plays the rôle of a wronskian for system $R$. Sup-

pose, however, that each of the $(r+1)$-rowed determinants vanishes somewhere in $A$. They can not all be identically zero in $A$, because then the wronskian of $y_1, y_2, \cdots, y_n$ would be zero also, contrary to hypothesis. Consequently one of these determinants, say that formed from the first $r+1$ columns of $M$, is different from zero at some point of $A$, viz., $(a_1, \cdots, a_p)$, and therefore, since it is continuous, it differs from zero in a $p$-dimensional region $A'$ containing this point. Then in this region $A'$, every solution of system $R$ is linearly dependent on $y_1, y_2, \cdots, y_{r+1}$, so that in $A'$ every solution of (7), being also a solution of $R$, will be linearly dependent on $y_1, y_2, \cdots, y_{r+1}$. Then in particular we shall have in $A'$

$$y_{r+2} = c_1 y_1 + c_2 y_2 + \cdots + c_{r+1} y_{r+1},$$

the $c$'s being constants not all zero. Therefore the function

$$y = y_{r+2} - c_1 y_1 - c_2 y_2 - \cdots - c_{r+1} y_{r+1}$$

will vanish identically in $A'$, and so will its derivatives $y^{(1)}, y^{(2)}, \cdots, y^{(n)}$. But it is easily seen that if system (7) has $n$ solutions $y_1, y_2, \cdots, y_n$ for which the wronskian vanishes nowhere in $A$, then any solution all of whose primary derivatives vanish at some point of $A$ will vanish identically in $A$. Consequently the function $y$ considered above vanishes identically not only in $A'$, but throughout $A$. Therefore we conclude that system (7) would have only $r+1$ linearly independent solutions, which contradicts the hypothesis.

We now proceed to show how system (7) may be reduced to a similar system for which the primary derivatives form a normal set, under the hypothesis that system (7) has exactly $n$ fundamental solutions $y_1, y_2, \cdots, y_n$ whose wronskian $W$ does not vanish at any point of a region $A$ of $p$ dimensions. We assume, then, that the $n$ primary derivatives of system (7) do not form a normal set. Then we may select from among these primary derivatives a normal set which consists of at least one element, namely $y$; let this normal set be $y, y^{(1)}, y^{(2)}, \cdots, y^{(r)}, (r < n-1)$. Suppose, furthermore, that it is the *largest* normal set which may be formed from among the primary derivatives. Then of the first derivatives of $y, y^{(1)}, y^{(2)}, \cdots, y^{(r)}$ none is a primary derivative of system (7), for otherwise our normal set could be enlarged.* Now, all of the first derivatives just mentioned are expressible linearly in terms of the primary derivatives of system (7), the coefficients in the expressions being continuous in $A$. This gives us $p(r+1)$ equations, in the right-hand members of which appear some or all of the quantities $y, y^{(1)}, y^{(2)}, \cdots,$

* Lest a misunderstanding occur here, we point out that if $y^{(i)}$ and $y^{(m)}$ are primary derivatives, and if $\partial y^{(i)}/\partial u_j = y^{(m)}$, then this equation is one of the equations of system (7); also, if $y^{(i)}$ be one of the normal set $y, y^{(1)}, \cdots, y^{(r)}$ under consideration, then $y^{(m)}$ belongs to the said normal set also.

$y^{(r)}$. But in the right-hand member of at least one of these equations must appear some one of the primary derivatives, say $y^{(s)}$, which is not one of our normal set, for otherwise the $p(r+1)$ equations would constitute a system of form (7) whose primary derivatives are $y, y^{(1)}, y^{(2)}, \cdots, y^{(r)}$, i. e., the said $p(r+1)$ equations would constitute a system $R$ such as is considered in the lemma proved above, and the original system (7) could then have no more than $r+1$ linearly independent solutions, contrary to hypothesis. Consequently, of the first derivatives of $y, y^{(1)}, y^{(2)}, \cdots, y^{(r)}$, at least one has a linear expression, in terms of the primary derivatives, which contains a primary derivative $y^{(s)}$ which is not a member of the normal set $y, y^{(1)}, y^{(2)}, \cdots, y^{(r)}$. In other words, the coefficient of $y^{(s)}$ in this expression is not identically zero; since this coefficient is continuous, there must exist a $p$-dimensional sub-region of $A$, say $A_1$, at no point of which it vanishes. Consequently, in the region $A_1$ we may express the primary derivative $y^{(s)}$ linearly in terms of all the other primary derivatives, and the additional quantity which occurs in the left-hand member of the equation from which we have just found the expression for $y^{(s)}$. This additional quantity, we recall, is one of the first derivatives of one of the elements of the normal set $y, y^{(1)}, y^{(2)}, \cdots, y^{(r)}$, and is not itself a primary derivative. Let us denote it by $y^{(t)}$. Then throughout (7) we may replace $y^{(s)}$ everywhere by its equivalent expression, and thus obtain a new system of differential equations, of form (7), in which the primary derivatives are the same except that $y^{(t)}$ has taken the place of $y^{(s)}$. This new set therefore contains in it a normal set composed of the previous normal set $y, y^{(1)}, y^{(2)}, \cdots, y^{(r)}$ and the additional derivative $y^{(t)}$. The new wronskian of the system of differential equations vanishes nowhere in the $p$-dimensional region $A_1$. But although the coefficients in system (7) are supposed to have all of their first derivatives, it is logically possible that all of the first derivatives do not exist for the coefficients of the new system. We shall suppose that these derivatives do exist.* Then on the new system of differential equations we may proceed as before, enlarging step by step the number of primary derivatives which form a normal set until finally a system is obtained for which all of the $n$ primary derivatives form a normal set. In this reduction we have supposed of course that certain derivatives exist for the coefficients in system (7). In any event, it is clear that such derivatives need not exist of order higher than $n-r-1$, where $r$ is the number which appears in the above discussion.

We shall henceforth suppose that the primary derivatives in system (7)

* The writer does not know whether they exist as a consequence of the previous hypotheses on system (7) or not. Since the form to which we shall ultimately reduce our system seems essential for the proof of the existence theorem, it would be of interest to have this point settled.

form a normal set, and shall then speak of system (7) as being in the *normal form*. The chief importance of the normal form lies in the fact that system (7) in that form may be regarded as a system of $np$ partial differential equations of the first order in the $p$ independent variables $u_1, u_2, \cdots, u_p$, and the $n$ dependent variables $y, y^{(1)}, y^{(2)}, \cdots, y^{(n-1)}$, since among equations (7) occur a certain number which define each primary derivative as the first derivative of another primary derivative belonging to the same chain.*

## 6. The canonical form of a completely integrable system

It is well known that by a transformation of the dependent variable of the form $y = \lambda(x)\bar{y}$ any ordinary homogeneous linear differential equation of the $n$th order,

$$\frac{d^n y}{dx^n} + p_1 \frac{d^{n-1} y}{dx^{n-1}} + \cdots + p_n y = 0,$$

whose coefficients are continuous functions of $x$, may be transformed into one of the same form in which the coefficient $p_1$ is zero. We shall now show that a similar thing may be done for certain completely integrable systems.

We introduce a new dependent variable, $\bar{y}$, connected with the old by the equation

$$y = \lambda\bar{y},$$

where $\lambda$ is a function of $u_1, u_2, \cdots, u_p$. Any derivative of $y$, say $y^{(j)}$, is connected with derivatives of the new dependent variable, $\bar{y}$, by an equation of the form

$$y^{(j)} = \lambda\bar{y}^{(j)} + \cdots,$$

where the part omitted in the right-hand member is linear in derivatives of $\bar{y}$, and contains all those derivatives of $\bar{y}$ from which $\bar{y}^{(j)}$ may be obtained by differentiation.

Let us now carry out this transformation on a particular kind of completely integrable system, namely, one in which the primary derivatives not only form a normal set, but have the additional property that, if a derivative $y^{(j)}$ belong to the set, then *every* derivative of lower order from which $y^{(j)}$ is obtainable by successive differentiation is likewise a primary derivative. It is not difficult to see that in this case, if $y$ be replaced everywhere by $\lambda\bar{y}$, the system of differential equations is changed into a new system, with the dependent variable $\bar{y}$, which new system may be thrown into exactly the same form as the original system. Let us write the new system

$$\frac{\partial \bar{y}^{(j)}}{\partial u_k} = \sum_{i=0}^{n-1} \bar{a}_i^{(j,k)} \bar{y}^{(i)} \tag{14}$$

$$(j = 0, 1, \cdots, n-1;\ k = 1, 2, \cdots, p),$$

* See footnote on page 503.

where the coefficients $\bar{a}_i^{(j,k)}$ are functions of the old coefficients $a_i^{(j,k)}$ (of system (7)) and also of $\lambda$ and its successive derivatives. Of course the actual expressions depend upon the form of each individual primary derivative, which form we have not specified.

We may, however, obtain explicit results with regard to the transformation of the wronskian and its first derivatives. Let us denote by $\overline{W}$ the wronskian for the transformed system (14). We recall that the set of primary derivatives has the property that if $y^{(j)}$ belong to the set all the derivatives of $y$ from which $y^{(j)}$ may be obtained by differentiation also belong to the set. Consequently the wronskian $\overline{W}$ for the new system is easily seen to be related to the wronskian for the old system by the equation

$$W = \lambda^n \overline{W}. \tag{15}$$

Differentiating this equation with respect to $u_k$, we have

$$\frac{\partial W}{\partial u_k} = \lambda^n \frac{\partial \overline{W}}{\partial u_k} + n\lambda^{n-1} \overline{W} \frac{\partial \lambda}{\partial u_k},$$

or, using equations (13) and (11),

$$\frac{\partial f}{\partial u_k} W = \lambda^n \overline{W} \frac{\partial \bar{f}}{\partial u_k} + n\lambda^{n-1} \overline{W} \frac{\partial \lambda}{\partial u_k},$$

where $\bar{f}$ denotes the function which takes the place of $f$ for the new system of differential equations (14). Substituting for $\overline{W}$ its value as obtained from (15), and recalling that $W$ is nowhere zero in the region $A$, we find that

$$\lambda \frac{\partial \bar{f}}{\partial u_k} = \lambda \frac{\partial f}{\partial u_k} - n \frac{\partial \lambda}{\partial u_k} \qquad (k = 1, 2, \cdots, p).$$

This result enables us to state the theorem:

THEOREM IX. *Suppose a system of partial differential equations of the form* (7) *has the following properties*:

1°. *It is completely integrable, with n fundamental solutions whose wronskian* (3) *vanishes at no point of A*.

2°. *The set of primary derivatives is such that, if $y^{(j)}$ be any one of the set, then all the derivatives of lower order from which $y^{(j)}$ may be obtained by differentiation also belong to the set.*

3°. *All of the primary derivatives exist for each of the np coefficients $a_j^{(j,k)}$ $(j = 0, 1, \cdots, n-1;\ k = 1, 2, \cdots, p)$.*

*Then the system of differential equations* (7) *may be transformed into another uniquely determined completely integrable system* (14) *for which all of the quantities*

$$\frac{\partial \bar{f}}{\partial u_k} \equiv \sum_{j=0}^{n-1} \bar{a}_j^{(j,k)} \qquad (k = 1, 2, \cdots, p),$$

*are zero, by the transformation of the dependent variable* $y = \lambda\bar{y}$, *where*

$$\lambda\frac{\partial f}{\partial u_k} - n\frac{\partial \lambda}{\partial u_k} = 0, \qquad \lambda = \text{const. } e^{f/n}. \tag{16}$$

The system (14) thus obtained will be unique, and its coefficients $\bar{a}_i^{(j,k)}$ will be explicit expressions in the coefficients of system (7) and their derivatives. Let us call this uniquely determined system the *canonical form* of system (7).

It should be noted that part 3° of the hypothesis is necessary in order that $\lambda\bar{y}$ may be substituted for $y$ throughout equation (7).

Let us return to the general transformation $y = \lambda\bar{y}$, which converts system (7) into system (14); we now suppose that system (14) is not in the canonical form. The coefficients of system (14) are expressible explicitly in terms of the coefficients of (7) and their derivatives, and $\lambda$ and its derivatives. Consider any function $\overline{F}$ of the coefficients of (14) and their derivatives. Let $F$ be the same function of the corresponding coefficients of (7) and their derivatives. The function $\overline{F}$ is equal identically to a function of $\lambda$ and the coefficients of (7), and the derivatives of these quantities. If in particular this latter function is related to $\overline{F}$ by the identity

$$F = \lambda^r \overline{F},$$

where $r$ may or may not be zero, we call the function $F$ a *seminvariant* of system (7).

The following theorem is a consequence of the fact that system (7) and the transformed system (14) have exactly the same canonical form:

THEOREM X. *The coefficients of the canonical form are seminvariants of system* (7). *All other seminvariants are functions of these and of their derivatives, if such derivatives exist.*

The reduction of a completely integrable system to the canonical form is seen by the above discussion to be a purely mechanical process, performable in any case coming under the hypotheses of Theorem IX. This fact is of practical importance in projective differential geometry, if the general method of Wilczynski be followed. We shall return to this question in the next section, applying the method to a particular geometric problem.

In § 8, we shall give an existence theorem for systems of form (7) in which the primary derivatives constitute a normal set. It should be noted that in the hypothesis 2° of Theorem IX, however, the set of primary derivatives is still further restricted; as a matter of fact, we may show by an example that this restriction is in general necessary, so that not every completely integrable system may be reduced to a unique canonical form. Denoting by subscripts partial differentiation with respect to the two independent

variables $u$ and $v$, let the primary derivatives be $y$, $y_u$, $y_{uv}$. These form a normal set. If $y_1$, $y_2$, $y_3$ be functions of $u$, $v$ for which in the region $A$ the wronskian

$$W = |y, y_u, y_{uv}|$$

is nowhere zero, we may set up the completely integrable system having $y_1$, $y_2$, $y_3$ as a fundamental set of solutions:

$$\begin{aligned} y_v &= ay + by_u + cy_{uv}, \\ y_{uu} &= a'y + b'y_u + c'y_{uv}, \\ y_{uuv} &= a''y + b''y_u + c''y_{uv}, \\ y_{uvv} &= a'''y + b'''y_u + c'''y_{uv}, \end{aligned}$$

where in general $c \neq 0$. Making the transformation $y = \lambda\bar{y}$, we find that $W$ is transformed into

$$W = |\lambda\bar{y}, \lambda\bar{y}_u + \lambda_u\bar{y}, \lambda\bar{y}_{uv} + \lambda_v\bar{y}_u + \lambda_u\bar{y}_v + \lambda_{uv}\bar{y}| = (\lambda^3 + c\lambda^2\lambda_u)\overline{W},$$

in which use has been made of the first of the system of differential equations. Now, if $c \neq 0$—a condition which may be realized at pleasure—this relation between $W$ and $\overline{W}$ is not of the form (15), so that the differential equation (16) which determines $\lambda$ is replaced by one of the second order. Therefore, although the system may be reduced to the canonical form, this form will not be unique.

Of course, instead of making the said hypothesis 2° of Theorem IX, we might demand merely that the wronskian and its transform be related by equation (15). Our example would then be included in this case, if $c$ were identically zero in $A$.

### 7. Geometric application; curvilinear coördinates*

Completely integrable systems of the kind we have been considering have attained increased importance of late owing to their systematic use in certain fields of projective differential geometry. Since any solution of the system is linearly dependent on any fundamental set of solutions, it follows that a geometric configuration defined by a fundamental set of solutions is a projective transformation of the configuration defined by any other fundamental set of solutions. Consequently, any geometric property expressed in terms of the completely integrable system, being common to all the projective transformations of the configuration defined by a fundamental set of solutions of the system, is a projective property. Wilczynski has developed a general method for dealing with certain questions from this point of view, and it

* This section is independent of § 8.

is in connection with this method that the results of § 6 find an important application.

Let us consider in particular the general subject of curvilinear coördinates in space of $n$ dimensions.* If $y_1, y_2, \cdots, y_{n+1}$ be interpreted as homogeneous coördinates of a point in this space, the equations

$$y_i = y_i(u_1, u_2, \cdots, u_n) \qquad (i = 1, 2, \cdots, n+1) \tag{17}$$

will define a system of curvilinear coördinates, provided the ratios of the $y$'s do not all reduce to functions of fewer than $n$ of the independent variables $u_1, u_2, \cdots, u_n$. We shall suppose that all derivatives of any order exist for the functions $y_i$. Since at least one of these functions, say $y_{n+1}$, is different from zero at some point, it will be different from zero in an $n$-dimensional region about this point. We wish, then, the analytic condition that the $n$ ratios $\eta_i \equiv y_i/y_{n+1}$, $(i = 1, 2, \cdots, n)$ actually define $\infty^n$ points. This condition is the non-vanishing of the jacobian of the functions $\eta_i$, and this may be reduced to the non-vanishing of the determinant

$$W = \begin{vmatrix} y_1, & y_2, & \cdots, & y_{n+1} \\ y_1^{(1)}, & y_2^{(1)}, & \cdots, & y_{n+1}^{(1)} \\ y_1^{(2)}, & y_2^{(2)}, & \cdots, & y_{n+1}^{(2)} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ y_1^{(n)}, & y_2^{(n)}, & \cdots, & y_{n+1}^{(n)} \end{vmatrix}, \tag{18}$$

where $y_1^{(j)}$, $y_2^{(j)}$, etc., denote $\partial y_1/\partial u_j$, $\partial y_2/\partial u_j$, etc. We may therefore set up a completely integrable system of form (7) of which the $n+1$ functions $y_i$ constitute a fundamental set of solutions, and in which the primary derivatives are $y, y^{(1)}, y^{(2)}, \cdots, y^{(n)}$. This system is

$$\frac{\partial y}{\partial u_k} = y^{(k)}, \qquad \frac{\partial y^{(j)}}{\partial u_k} = \sum_{i=0}^{n} a_i^{(j,k)} y^{(i)} \qquad (j, k = 1, 2, \cdots, n). \tag{19}$$

Let us differentiate $W$ with respect to $u_k$; then on using equations (19) we find without difficulty that

$$\frac{\partial W}{\partial u_k} = (a_1^{(1,k)} + a_2^{(2,k)} + \cdots + a_n^{(n,k)})\, W.$$

We may therefore put

$$\sum_{j=1}^{n} a_j^{(j,k)} \equiv \frac{\partial f}{\partial u_k} \qquad (k = 1, 2, \cdots, n).$$

* The case $n = 2$, i. e., nets of plane curves, has been studied by Wilczynski, *One-parameter families and nets of plane curves*, these Transactions, vol. 12 (1911), pp. 473–510. The present writer considered the case $n = 3$ in his Columbia dissertation, *Projective differential geometry of triple systems of surfaces*, Lancaster, Pa., Press of The New Era Printing Company, 1913.

Now, there is an arbitrary feature introduced in equations (17) because of the homogeneity of the coördinates. If $\lambda$ be any function of $u_1, u_2, \cdots, u_n$, then the points defined by the coördinates $\bar{y}_i$ $(i = 1, 2, \cdots, n+1)$ will be the same as those defined by equations (17), if $y_i = \lambda \bar{y}_i$. Consequently, if a projective property of the configuration is to be expressed in terms of the coefficients of the completely integrable system (19), it must be in terms of what we called in § 6 the *seminvariants* of the system, viz., those combinations of the coefficients of equations (19) and their derivatives which remain essentially unchanged when the system is subjected to an arbitrary transformation of the form

$$y = \lambda \bar{y}.$$

Let us carry out this transformation of the dependent variable, thus converting system (19) into one of the same form,

$$\frac{\partial \bar{y}}{\partial u_k} = \bar{y}^{(k)}, \qquad \frac{\partial \bar{y}^{(j)}}{\partial u_k} = \sum_{i=0}^{n} \bar{a}_i^{(j,k)} \bar{y}^{(i)} \qquad (j, k = 1, 2, \cdots, n). \tag{20}$$

The coefficients of this system are functions of the coefficients of (19) and their derivatives, and also of $\lambda$ and its derivatives. In particular,

$$\bar{a}_j^{(j,k)} = a_j^{(j,k)} - \lambda^{(k)}/\lambda \quad (j \neq k), \qquad \bar{a}_j^{(j,j)} = a_j^{(j,j)} - 2\lambda^{(k)}/\lambda \quad (j, k = 1, 2, \cdots, n),$$

so that, on denoting by $\partial \bar{f}/\partial u_k$ the quantity $\sum_{j=1}^{n} \bar{a}_j^{(j,k)}$, we find directly that

$$\lambda \frac{\partial \bar{f}}{\partial u_k} = \lambda \frac{\partial f}{\partial u_k} - (n+1) \frac{\partial \lambda}{\partial u_k} \qquad (k = 1, 2, \cdots, n).$$

This also follows from the general results of § 6. Consequently, the quantities $\bar{f}^{(k)} \equiv \partial \bar{f}/\partial u_k$ will all be zero if $\lambda$ be chosen as a solution of

$$\lambda f^{(k)} - (n+1)\lambda^{(k)} = 0 \qquad (k = 1, 2, \cdots, n),$$

which gives

$$\lambda = \text{const.}\ e^{f/(n+1)}.$$

With this choice of $\lambda$ the coefficients of the new system (20) have the values

$$\begin{aligned}
&\bar{a}_0^{(j,j)} = a_0^{(j,j)} + \frac{1}{n+1}\left( \sum_{i=1}^{n} a_i^{(j,j)} f^{(i)} - \frac{\partial f^{(j)}}{\partial u_j} - \frac{1}{n+1} f^{(j)} f^{(j)} \right), \\
&\bar{a}_i^{(j,j)} = a_i^{(j,j)} \ (i \neq j), \qquad \bar{a}_j^{(j,j)} = a_j^{(j,j)} - 2f^{(j)}, \\
&\bar{a}_0^{(j,k)} = a_0^{(j,k)} + \frac{1}{n+1}\left( \sum_{i=1}^{n} a_i^{(j,k)} f^{(i)} - \frac{\partial f^{(j)}}{\partial u_k} - \frac{1}{n+1} f^{(j)} f^{(k)} \right), \\
&\bar{a}_i^{(j,k)} = a_i^{(j,k)} \ (i \neq j, k), \qquad \bar{a}_j^{(j,k)} = a_j^{(j,k)} - f^{(k)}, \\
&\bar{a}_k^{(j,k)} = a_k^{(j,k)} - f^{(j)} \qquad (j \neq k;\ i, j, k = 1, 2, \cdots, n),
\end{aligned} \tag{21}$$

where throughout single superscripts denote differentiation. The quantities (21) are the fundamental seminvariants of system (19), and any seminvariant of the system is a function of these and their derivatives.

There still remains, however, an arbitrary feature in equations (17). This is in the choice of parameters. Since the system of curvilinear coördinates remains unchanged under the transformation

$$\bar{u}_k = U_k(u_k), \qquad (k = 1, 2, \cdots, n), \tag{22}$$

it follows that projective properties of the configuration are expressible in terms of those seminvariants which remain essentially invariant when system (19) is subjected to the transformation (22). It is possible to calculate explicitly, and without any difficulty, these *invariants*, but our object here is not to give a systematic theory of curvilinear coördinates.

The fact that even in this very general problem the invariants may be calculated explicitly, makes it highly probable that the analytic methods of Wilczynski are practicable in handling problems in the projective differential geometry, not only of configurations in spaces of a few dimensions, but also in space of $n$ dimensions. The only methods which have hitherto been used systematically in studying configurations in hyperspace are those of Segre.* The ideal method seems to the present writer to be one in which all the projective differential properties are expressible by means of a unique analytic apparatus, and this is to be found in the completely integrable system of which a fundamental system of solutions defines the particular geometric configuration under discussion.

## 8. The existence theorem for completely integrable systems. The linear dependence of solutions

The theorems of previous sections concerning the system of partial differential equations (7) presuppose the existence of a fundamental system of $n$ solutions. The question of this existence presents no essentially new difficulty, for the remark made at the end of § 5 renders immediately applicable to system (7), which is supposed to be in the normal form, a well-known existence theorem concerning systems of total differential equations.†

* For references to some of this work, see E. Bompiani, *Recenti progressi nella geometria proiettiva differenziale degli iperspazi*, Proceedings of the fifth international congress of mathematicians, Cambridge, 1912, vol. II, pp. 22–27.

In this connection should be mentioned also Prof. Wilczynski's Cleveland address, *Some general aspects of modern geometry*, Bulletin of the American Mathematical Society, vol. 19 (1913–14), pp. 331–342.

† Cf., for example, Ch.-J. de la Vallée Poussin, *Cours d'Analyse Infinitésimale*, 2me éd., vol. 2, chap. VII, § 6. The theorem is essentially that of Mayer, given in the memoir cited in § 3.

A simultaneous system of $n$ total differential equations in $n$ dependent variables $x_h$ $(h = 1, 2, \cdots, n)$ and $m$ independent variables $y_i$ $(i = 1, 2, \cdots, m)$ is of the form

$$dx_h = \sum_{i=1}^{m} a_h^{(i)}\, dy_i \qquad (h = 1, 2, \cdots, n;\ i = 1, 2, \cdots, m), \tag{23}$$

the $a$'s being continuous functions of the $x$'s and $y$'s in a region $A$. Let us write

$$\frac{\delta}{\delta y_k} = \frac{\partial}{\partial y_k} + \sum_{h=1}^{n} a_h^{(k)} \frac{\partial}{\partial x_h}.$$

Then we shall say that system (23) is *completely integrable* if its coefficients satisfy the identities

$$\frac{\delta a_h^{(i)}}{\delta y_k} = \frac{\delta a_h^{(k)}}{\delta y_i} \qquad (i, k = 1, 2, \cdots, m;\ h = 1, 2, \cdots, n). \tag{24}$$

We shall now state the existence theorem for this system.

THEOREM. *Suppose that the identities* (24) *subsist for system* (23) *throughout the region* $A$, *and let* $y_1^{(0)}, \cdots, y_m^{(0)}$; $x_1^{(0)}, \cdots, x_n^{(0)}$ *be a set of values of the independent and dependent variables in the interior of this region. Then there exists one and only one set of* $n$ *functions* $x_1, \cdots, x_n$ *of the* $m$ *independent variables* $y_1, \cdots, y_m$, *which satisfy the system of differential equations and which respectively take on the values* $x_1^{(0)}, \cdots, x_n^{(0)}$ *when* $y_1 = y_1^{(0)}, \cdots, y_m = y_m^{(0)}$.

Consider now system (7), viz.,

$$\frac{\partial y^{(j)}}{\partial u_k} = \sum_{i=0}^{n-1} a_i^{(j,k)}\, y^{(i)} \qquad (j = 0, 1, \cdots, n-1;\ k = 1, 2, \cdots, p),$$

in which we suppose that there are exactly $n$ primary derivatives, and that they form a normal set. Let us regard $y, y^{(1)}, \cdots, y^{(n-1)}$ as dependent variables; then the system becomes a very particular case of the system (23) which we have just been considering. The integrability conditions (24) take precisely the form of equations (9). Moreover, from the remark made at the end of § 5, system (7) when regarded as a system of partial differential equations of the first order with the dependent variables $y, y^{(1)}, \cdots, y^{(n-1)}$ is exactly equivalent to system (7) with its original significance, viz., that in which $y, y^{(1)}, \cdots, y^{(n-1)}$ are primary derivatives of the function $y$, each superscript denoting a certain particular, though unspecified, partial derivative of $y$. The exact equivalence of the two points of view follows from the fact that the primary derivatives form a normal set. In other words, if system (7) be regarded as a system of partial differential equations of the

first order in $n$ dependent variables, then a set of solutions $y, y^{(1)}, \cdots, y^{(n-1)}$ will have the property that $y^{(1)}, y^{(2)}, \cdots, y^{(n-1)}$ are exactly those derivatives of $y$ which the original superscript notation signified; and all this without requiring the addition of new equations to system (7).

The force of these remarks, especially the last, may be seen by means of the completely integrable system in two independent variables $u$ and $v$, with primary derivatives $y, y_u, y_v$:

$$(a) \qquad \begin{aligned} y_{uu} &= a y_u + b y_v + c y, \\ y_{uv} &= a' y_u + b' y_v + c' y, \\ y_{vv} &= a'' y_u + b'' y_v + c'' y, \end{aligned}$$

where the subscripts denote differentiation. This system may be replaced by a system with three dependent variables by means of the substitution

$$(b) \qquad y_u = y^{(1)}, \qquad y_v = y^{(2)},$$

and will then have the form

$$(a') \qquad \begin{aligned} y_u^{(1)} &= a y^{(1)} + b y^{(2)} + c y, \\ y_v^{(1)} &= a' y^{(1)} + b' y^{(2)} + c' y, \\ y_u^{(2)} &= a' y^{(1)} + b' y^{(2)} + c' y, \\ y_v^{(2)} &= a'' y^{(1)} + b'' y^{(2)} + c'' y. \end{aligned}$$

This is not in the form (7), however, because all of the first derivatives of all of the primary derivatives do not appear in the left-hand members of system $(a')$. But equations $(a')$ and $(b)$ together constitute a system of form (7). System $(a')$ alone is not equivalent to system $(a)$, but systems $(a')$ and $(b')$ together are equivalent to system $(a)$.

A better illustration is afforded by the system of form (7) introduced at the beginning of § 5, viz., one in which the primary derivatives are $y$ and $y_{uu}$, and which we there saw had no solution different from zero. If we put

$$y_{uu} = y^{(1)},$$

then the system in form (7) is

$$\begin{aligned} y_u &= a y + b y^{(1)}, & y_v &= a' y + b' y^{(1)}, \\ y_u^{(1)} &= a'' y + b'' y^{(1)}, & y_v^{(1)} &= a''' y + b''' y^{(1)}. \end{aligned}$$

These four equations, regarded as equations of the first order in the dependent variables $y, y^{(1)}$, will have sets of solutions $y, y^{(1)}$ different from zero if the integrability conditions (9) are satisfied for the system. But the function $y^{(1)}$ will not in general be the proper derivative of the function $y$, viz., $y_{uu}$. This

can not happen in the case of the other example, because equations ($b$), which form part of the system in the form (7), make it certain that a solution $y$, $y^{(1)}$, $y^{(2)}$ of the system consisting of equations ($a'$) and ($b$) will have the property that $y^{(1)}$ and $y^{(2)}$ will be the proper derivatives, $y_u$ and $y_v$ respectively, of $y$.

Let us, then, regard system (7) as a system of partial differential equations of the first order in the $n$ dependent variables $y$, $y^{(1)}$, $\cdots$, $y^{(n-1)}$. Applying the existence theorem for system (23), of which (7) is a particular case, we may state the following existence theorem:

THEOREM XI. *Suppose that in the system of partial differential equations*

$$\frac{\partial y^{(j)}}{\partial u_k} = \sum_{i=0}^{n-1} a_i^{(j,k)} y^{(i)} \qquad (j = 0, 1, \cdots, n-1;\ k = 1, 2, \cdots, p) \tag{7}$$

*the primary derivatives $y$, $y^{(1)}$, $\cdots$, $y^{(n-1)}$ form a normal set. Suppose further that in the region $A$ the coefficients $a_i^{(j,k)}$, which are functions of the $p$ independent variables $u_1, u_2, \cdots, u_p$, satisfy identically the integrability conditions*

$$\frac{\partial a_\nu^{(j,k)}}{\partial u_l} + \sum_{i=0}^{n-1} a_i^{(j,k)} a_\nu^{(i,l)} = \frac{\partial a_\nu^{(j,l)}}{\partial u_k} + \sum_{i=0}^{n-1} a_i^{(j,l)} a_\nu^{(i,k)} \qquad (\nu, j = 0, 1, \cdots, n-1;\ k, l = 1, 2, \cdots, p). \tag{9}$$

*Let $(u_1^{(0)}, u_2^{(0)}, \cdots, u_p^{(0)})$ be any point of $A$, and $y_0, y_0^{(1)}, \cdots, y_0^{(n-1)}$ be any set of $n$ constants. Then there exists one and only one function $y$ of the variables $u_1, u_2, \cdots, u_p$ which satisfies the system of differential equations, and whose primary derivatives $y$, $y^{(1)}$, $\cdots$, $y^{(n-1)}$ take on respectively the preassigned constant values $y_0, y_0^{(1)}, \cdots, y_0^{(n-1)}$ at the point $(u_1^{(0)}, u_2^{(0)}, \cdots, u_p^{(0)})$.*

As in the familiar case of ordinary differential equations, this theorem allows us to infer the existence of $n$ linearly independent solutions whose wronskian vanishes nowhere in $A$. In fact, we may construct a particular set of solutions having this property, by choosing solutions $Y_1, Y_2, \cdots, Y_n$ such that the primary derivatives of the solution $Y_i$ take on the values

$$Y_i^{(j)} = 0 \ (j \neq i-1), \qquad Y_i^{(j)} = 1 \ (j = i-1)$$

at the point $(u_1^{(0)}, u_2^{(0)}, \cdots, u_p^{(0)})$. At this point the wronskian of $Y_1, Y_2, \cdots, Y_n$ is then

$$\begin{vmatrix} 1, & 0, & \cdots, & 0 \\ 0, & 1, & \cdots, & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0, & 0, & \cdots, & 1 \end{vmatrix} = 1,$$

so that the wronskian vanishes nowhere in $A$, by Theorem VIII. Moreover, that solution of the system for which the primary derivatives take on the values $y_0, y_0^{(1)}, \cdots, y_0^{(n-1)}$ at the point $(u_1, u_2, \cdots, u_p)$ is given by

$$y = y_0 Y_1 + y_0^{(1)} Y_2 + y_0^{(2)} Y_3 + \cdots + y_0^{(n-1)} Y_n .$$

If a solution has all of its primary derivatives zero at any point of $A$, it therefore vanishes identically in $A$. Let $y_1, y_2, \cdots, y_n$ be any $n$ solutions of the completely integrable system; then no function, other than zero, of the form

$$y = c_1 y_1 + c_2 y_2 + \cdots + c_n y_n ,$$

can vanish together with its derivatives $y^{(1)}, y^{(2)}, \cdots, y^{(n-1)}$ at any point of $A$, unless the constants $c_i$ are all zero. This set of functions $y_1, y_2, \cdots, y_n$ has therefore the properties of the $n$ functions of Theorem III. We may therefore apply this theorem to prove the following one, which is also a consequence of the existence of $n$ solutions whose wronskian vanishes nowhere in $A$:

THEOREM XII. *If $y_1, y_2, \cdots, y_n$ be any $n$ solutions of the completely integrable system of Theorem XI, then the vanishing, anywhere in $A$, of the wronskian of these $n$ functions is a sufficient condition for their linear dependence.*

This theorem is included in the following more general one. They are both generalizations of theorems given by Bôcher for ordinary differential equations.*

THEOREM XIII. *If $y_1, y_2, \cdots, y_k$ $(k \leqq n)$ be solutions of the completely integrable system of Theorem XI, a sufficient condition that they be linearly dependent in $A$ is the identical vanishing in $A$ of all those $k$-rowed determinants of the matrix*

$$M \equiv M_{n-1}(y_1, y_2, \cdots, y_k)$$

*which contain elements of the first row of the matrix.*

The proof of this theorem follows immediately from Theorem IV. The matrix $M$ of Theorem IV is in the present case the matrix formed from the primary derivatives of the functions $y_1, y_2, \cdots, y_k$, while the augmented matrix $M'$ has additional rows consisting of the first derivatives of the quantities appearing in $M$. But these first derivatives are linear combinations of the primary derivatives, in virtue of the completely integrable system, so that if any $\nu$ columns $(\nu \leqq k)$ be selected from $M$, the vanishing of all the $\nu$-rowed determinants formed from these columns and containing elements of the first row of $M$ carries with it the vanishing of all the $\nu$-rowed determinants formed from the same $\nu$ columns of $M'$ and containing elements of the first row of $M$. Moreover, no solution of the completely integrable system, other than zero, can have all of its primary derivatives vanish simultaneously at any point of $A$. Consequently all of the hypotheses of Theorem IV are satisfied by our functions $y_1, y_2, \cdots, y_k$, and the present theorem follows at once.

* M. Bôcher, loc. cit., these Transactions, vol. 2 (1901), pp. 139–149, theorem VII.

Enough has been given to show the marked analogy between ordinary homogeneous linear differential equations of the $n$th order and the completely integrable systems which have come under discussion in this paper. This analogy indicates the direction in which facts concerning ordinary differential equations may be extended to systems of partial differential equations. It is obvious how an adjoint system may be defined; the question naturally presents itself as to whether boundary value problems similar to those for ordinary differential equations may arise in connection with completely integrable systems. Questions concerning analytic functions of several variables suggest themselves also. In particular, can the solutions of a given completely integrable system be characterized by their properties as analytic functions, that is, does there exist a solution of the Riemann problem for such a system? Unfortunately, the theory of functions of several variables has not yet reached a degree of development comparable with that of functions of one variable, so that it is doubtful whether the generalization is feasible at present. However, all of these subjects certainly deserve investigation.

HARVARD UNIVERSITY,
*November*, 1915

# ON THE FACTORIZATION OF EXPRESSIONS OF VARIOUS TYPES*

BY

HENRY BLUMBERG

## INTRODUCTION

Not a few mathematicians have dealt with the problem of setting up criteria by means of which the irreducibility of certain expressions in certain domains may be seen at a glance from the character of the expressions. Gauss, Kronecker, Schoenemann, Eisenstein, Dedekind, Floquet, Koenigsberger, Netto, Perron, M. Bauer, and Dumas have written on the subject.† A simple example of such criteria is the well-known Schoenemann-Eisenstein theorem, which is essentially as follows:‡ If the polynomial

$$C = y^n + c_1 y^{n-1} + \cdots + c_n,$$

with integral coefficients, is such that a fixed prime $p$ goes into every

$$c_\nu \{\nu = 1, 2, \cdots, n\},$$

but $p^2$ does not go into $c_n$, then $C$ is irreducible in the domain of rational numbers. This theorem, while comparatively not rich in content, may be regarded as the nucleus of the work of the authors mentioned.

With the exception of Floquet and Koenigsberger, the authors mentioned deal exclusively with the case where the expressions are polynomials and chiefly with the case of polynomials whose coefficients belong to the domain

* Presented to the Society in less general form, November 28 and December 29, 1914.

† Gauss, *Disquisitiones arithmeticae* (1801), Art. 341; Kronecker, Journal für Mathematik, vol. 29 (1845), p. 280, vol. 100 (1887), p. 79, Journal de Mathématiques, vol. 19 (1854), p. 177, ser. 2, vol. 1 (1856), p. 399; Schoenemann, Journal für Mathematik, vol. 32 (1846), p. 100, vol. 40 (1850), p. 188; Eisenstein, Journal für Mathematik, vol. 39 (1850), p. 160; Dedekind, Journal für Mathematik, vol. 54 (1857), p. 27; Floquet, Annales de l'école normale (1879) [cited by Koenigsberger]; Koenigsberger, Journal für Mathematik, vol. 115 (1895), p. 53, vol. 121 (1900), p. 320, Mathematische Annalen, vol. 53 (1900), p. 49; Netto, Mathematische Annalen, vol. 48 (1897), p. 81 and *Algebra* (1896), vol. 1, lecture 5; Perron, Mathematische Annalen, vol. 60 (1905), p. 448, Journal für Mathematik, vol. 132 (1907), p. 288; M. Bauer, Journal für Mathematik, vol. 128 (1905), p. 87 and p. 298, vol. 132 (1907), p. 21, vol. 134 (1908), p. 15; Dumas, Journal de Mathématiques, ser. 6, vol. 2 (1906), p. 191.

‡ Schoenemann and Eisenstein, l. c. This theorem has been wrongly attributed to Eisenstein alone by most of the mathematicians who have had occasion to refer to it.

of rational numbers. Floquet and Koenigsberger have extended the investigation to linear homogeneous differential expressions.

A variety of methods have been employed. Thus the theory of algebraic fields, the theory of algebraic functions and the character of the solutions of linear homogeneous differential equations have been used.* Elementary methods, requiring no such means and based on the study of the expression as such without regard to the character of the solutions of the equation obtained by setting the expression equal to zero, have mostly succeeded in yielding only the less general results.†

The principal aim of the present paper is to communicate a general theorem and certain immediate consequences of it, which include as special cases, with minor exceptions,‡ all the results heretofore obtained in the indicated field of investigation. This theorem is given in § 6, and for convenience of designation referred to as the "Theorem on the Product"; its direct consequences are the "Theorem on Common Divisors" (§ 6), the "Factorization Theorem" and the "Criterion for Irreducibility" (§ 7). Thus the work on the subject before the publication of the Schoenemann-Eisenstein theorem is summarized and generalized by that theorem, which is an exceedingly special case of the "Criterion for Irreducibility." Moreover, the Floquet theorem, all the results in the first and the third papers of Koenigsberger, the chief results in his second paper, all the results of Netto, all the results in the 1905 paper of Perron, the chief results in his 1907 paper and the chief results of Bauer and of Dumas come under this "Theorem on the Product."§

A former note‖ was intended to give a summary of the present paper. Since then, however, the author has found how to unify and to generalize his previous results by means of the present theory. The note materially extends most of the known results in the field under discussion. For further details, the reader is referred to the note itself. Its relation to the present paper may be approximately described as follows: While the "Theorem on the Product" asserts that *every* pair $(m_\zeta, t_\zeta)$, $\{\zeta = 1, 2, \cdots, h\}$, of the characteristic set $S_C$ satisfies the condition imposed upon it by the relation $S_A + S_B = S_C$, the note makes use merely of the fact that the *first* pair satisfies this condition.¶ Attention is also directed to the replacing of the *three* theories of the note by the *single* theory $\mathfrak{T}_\epsilon$ of the present paper.

* Heretofore the results have not been carried over into the field of difference equations, possibly because of the recency of the publication of existence theorems in that field.

† Dumas's memoir, l. c., is a notable exception.

‡ For example, theorem I of Perron's 1907 paper, l. c., p. 291.

§ See previous citations.

‖ Proceedings of the National Academy of Sciences, June, 1915, pp. 374–381.

¶ See sections 4, 5, and 6.

The large variety of cases covered necessitates an abstract, postulational treatment. The different special theories (for polynomials with various kinds of coefficients and for difference and differential expressions with various kinds of coefficients) are thus made to appear as instances of an abstract general theory $\mathfrak{T}_\epsilon$ (see §§ 3 and 8). Only in such fashion can the interconnection of the more special results be brought into clear light and their unification effected. This abstract treatment is, moreover, especially fitting here, because a small number of simple assumptions is sufficient for the foundations of the theory.

The "Theorem on the Product" applies to Situations 4, 5 (for $k > 1$), 6, 7, 8, 9 (§ 3), not treated before, as far as the author knows, and enriches in a marked degree the results for Situations 3 (§ 3) and 10 (§ 8) already studied. Nevertheless, it does not go beyond the results of Dumas for the case of polynomials whose coefficients are rational numbers or Hensel $p$-adic numbers. Our general theory may, in fact, be regarded as an extension to an abstract situation of Dumas's theorem, l. c., p. 217.*

The remarks at the close of § 7 indicate the great difficulty in obtaining a more exhaustive criterion for irreducibility than the one actually obtained.

A sharp distinction has been made in the literature (as regards polynomials) between those investigations that are based on the *divisibility* properties of the coefficients and those that proceed from the consideration of the *magnitude* of the coefficients.† One of the interesting results is that the gap between these two types of investigation is bridged, both appearing subordinate to our general theory.

It is easy to see, as Koenigsberger has pointed out by an example, that the Schoenemann-Eisenstein theorem cannot be directly extended to the case of linear homogeneous differential expressions. Our abstract treatment furthermore lays bare the underlying reason—by no means evident otherwise—why it breaks down, and at once indicates what analogous theorem may replace it. The "Criterion for Irreducibility" may, in fact, be properly regarded as a highly generalized Schoenemann-Eisenstein theorem for the

* It is interesting in this connection to compare several remarks in Perron's 1907 paper, l. c. Referring to the work of Dumas on polynomials with rational coefficients as related to that of his predecessors, Perron says (p. 288): "Zu *etwas* allgemeineren Resultaten ist neuerdings Herr G. Dumas gelangt." Later on in the same paper (p. 304), Perron designates his most general result, Theorem VIII, as "ein *sehr* allgemeines Kriterium." The italics are in both cases my own. It turns out, however, that Perron's criterion is a special case of our theory, which is a natural and easy generalization of Dumas's results.

† Cf. Perron, 1907, l. c., p. 288 and Loewy in Pascal's *Repertorium* (1910), *Analysis* II, pp. 292 and 293. The same kind of distinction may, of course, be made with regard to differential and other expressions.

case of linear homogeneous differential expressions with coefficients taken as in Situation 6, 7, or 10.*

Here is a brief summary of the contents: Section 1 shows by means of an illustrative example how the Theory $\mathfrak{T}_{-1}$ of Section 2 arises. Section 3 describes various concrete situations falling under the Theory $\mathfrak{T}_{-1}$. Section 4 defines "characteristic set"—a fundamental notion for our purposes—and Section 5, the "sum of two characteristic sets." Section 6 gives the proof of the "Theorem on the Product" by the aid of two lemmas (established in §§ 4 and 5). In Section 7, the "Factorization Theorem" and the "Criterion for Irreducibility" are obtained. Section 8 introduces the abstract theory $\mathfrak{T}_\epsilon$ for every real $\epsilon$ and for $\epsilon = -\infty$. The appendix gives concrete examples of our results, at the same time exhibiting many instances of contact with the literature.

To secure greater clearness, nearly all the letters used are made to have a fixed meaning throughout the paper. Thus, for example, throughout the paper, $A$, $B$, $C$ stand respectively for $(a_0, a_1, \cdots, a_r)$, $(b_0, b_1, \cdots, b_s)$, $(c_0, c_1, \cdots, c_n)$; $\alpha_\lambda$, $\beta_\mu$, $\gamma_\nu$, respectively for the "rank" of $a_\lambda$, $b_\mu$, $c_\nu$ (see § 2); $\{(m_1, t_1), (m_2, t_2), \cdots, (m_h, t_h)\}$, for the "characteristic set" $S_C$ of $C$ (see § 4); etc. Slight departures from this state of affairs will be evident from the context.

## 1. An illustrative example suggesting the abstract treatment

Instead of proceeding directly to the exposition of the abstract theory mentioned in the introduction, we shall foreshadow the abstract situation by drawing certain pertinent conclusions in regard to one of its concrete instances.† In this way, we believe, the assumptions used in § 2 will be sufficiently motivated for the reader.

Let $\mathfrak{S}$ denote the class of rational functions of $x$ with complex numerical coefficients; and let

$$C = c_0(x)\frac{d^n y}{dx^n} + c_1(x)\frac{d^{n-1} y}{dx^{n-1}} + \cdots + c_n(x)y$$

be a linear homogeneous differential expression whose coefficients belong to $\mathfrak{S}$. We are concerned with the representation of $C$ as a product $A \cdot B$, where

$$A = a_0(x)\frac{d^r y}{dx^r} + a_1(x)\frac{d^{r-1} y}{dx^{r-1}} + \cdots + a_r(x)y$$

* Curiously, Koenigsberger himself has obtained results (essentially for Situation 10, § 8) that may be properly regarded as extensions of the Schoenemann-Eisenstein theorem to differential expressions. Lacking the abstract background, he failed, however, to notice this relation.

† The instance given has actually been the starting point of the author's present investigations.

and

$$B = b_0(x)\frac{d^s y}{dx^s} + b_1(x)\frac{d^{s-1} y}{dx^{s-1}} + \cdots + b_s(x)y,$$

the coefficients $a_\lambda$ and $b_\mu$ belonging to $\mathfrak{S}$. As usual in such cases, $A \cdot B$ means here the *symbolic* product of the two differential expressions, so that

$$A \cdot B = a_0\frac{d^r B}{dx^r} + a_1\frac{d^{r-1} B}{dx^{r-1}} + \cdots + a_r B$$

$$= a_0 b_0\frac{d^{r+s} y}{dx^{r+s}} + \left(ra_0\frac{db_0}{dx} + a_1 b_0 + a_0 b_1\right)\frac{d^{r+s-1} y}{dx^{r+s-1}} + \cdots$$

$$= \sum_{\nu=0}^{r+s=n} c_\nu \frac{d^{n-\nu} y}{dx^{n-\nu}},$$

where

$$c_\nu = \sum_{\mu=0}^{\nu}\sum_{\lambda=0}^{\nu-\mu}\binom{r-\lambda}{\nu-\lambda-\mu} a_\lambda \frac{d^{\nu-\mu-\lambda} b_\mu}{dx^{\nu-\mu-\lambda}},$$

$$\binom{r-\lambda}{\nu-\lambda-\mu} = \frac{(r-\lambda)!}{(\nu-\lambda-\mu)!(r-\nu+\mu)!}.$$

Understanding, as usual, by the degree of a rational fraction the degree of its numerator minus the degree of its denominator, we desire to call attention to several evident relations between the degrees of the $c$'s on the one hand and the degrees of the $a$'s and of the $b$'s on the other. From the expression of $c_\nu$ in terms of $a_\lambda$ and $b_\mu$, we see that $c_\nu$ is a sum of products of the form $h_{\lambda\mu\sigma}\, a_\lambda\, d^\sigma b_\mu/dx^\sigma$, where $\lambda + \mu + \sigma = \nu$ and $h_{\lambda\mu\sigma}$ is a numerical coefficient. We now use the following facts regarding the degrees of products, sums, and derivatives of rational fractions: (1) The degree of a product is equal to the sum of the degrees of the factors. (2) The degree of a sum is at most equal to the highest degree $d$ attained by its terms, and exactly equal to $d$ if only one term has the degree $d$. (3) The degree of $df(x)/dx$, $f(x)$ being a rational fraction, is at most one less than the degree of $f(x)$. To make these laws hold without exception, *we agree to set the degree of* 0 *equal to* $-\infty$ and to calculate with $-\infty$ in the usual manner. Thus $-\infty \pm 1 = -\infty$, $-\infty + (-\infty) = -\infty$, etc. Let $\alpha_\nu$, $\beta_\nu$, $\gamma_\nu$ denote the degrees of $a_\nu$, $b_\nu$, $c_\nu$. It is then clear that $\gamma_\nu$ is at most equal to the largest value attained by the numbers of the set $\alpha_\lambda + \beta_\mu - \sigma$, when $\lambda$, $\mu$, and $\sigma$ vary so as to satisfy the relation $\lambda + \mu + \sigma = \nu$. This fact we express conveniently as follows:*

* The fact that $\sigma \leqq r - \lambda$ really imposes more stringent conditions upon $\gamma_\nu$ than we use. Because of our weaker conditions, we gain not only generality but also the important advantage of symmetry, by virtue of which later proofs are materially shortened. Condition (*b*) below is also weaker than the equivalent of the second part of (2) above (regarding the degree of a sum).

(*a*) $\gamma_\nu \leqq \max_{\lambda+\mu+\sigma=\nu} (\alpha_\lambda + \beta_\mu - \sigma)$, $\{\nu = 0, 1, 2, \cdots, n\}$.

Furthermore, suppose that for the fixed values $\lambda = \bar{\lambda}$, $\mu = \bar{\mu}$, $\sigma = 0$ (where $\bar{\lambda} + \bar{\mu} = \nu$), $\alpha_{\bar{\lambda}} + \beta_{\bar{\mu}}$ is greater than $\alpha_\lambda + \beta_\mu - \sigma$ for every $[\lambda, \mu, \sigma] \neq [\bar{\lambda}, \bar{\mu}, 0]$ satisfying the relation $\lambda + \mu + \sigma = \nu$; it follows that $\gamma_\nu = \alpha_{\bar{\lambda}} + \beta_{\bar{\mu}}$. Hence, in addition to (*a*), we have:

(*b*) $\gamma_\nu = \alpha_{\bar{\lambda}} + \beta_{\bar{\mu}}$ $\{\bar{\lambda} + \mu = \nu\}$ if $\alpha_{\bar{\lambda}} + \beta_{\bar{\mu}} > \alpha_\lambda + \beta_\mu - \sigma$ $\{\lambda + \mu + \sigma = \nu\}$ for every $[\lambda, \mu, \sigma] \neq [\bar{\lambda}, \bar{\mu}, 0]$, $\{\nu = 0, 1, 2, \cdots, n\}$.

It will appear in the sequel that the simple and evident properties (*a*) and (*b*) of linear homogeneous differential expressions with rational coefficients essentially suffice for the demonstration of our factorization results concerning such expressions. But the properties (*a*) and (*b*) are shared by many other expressions (see the instances in § 3). To dispense with the duplication of proofs and for other reasons given in the introduction, we proceed (in the next section) to the exposition of an abstract theory $\mathfrak{T}_{-1}$, of which the illustrative example described in this section is a special instance.* The domain $\mathfrak{S}$ of rational functions will be replaced by an unrestricted aggregate $\mathfrak{S}$; the differential expressions, by "parentheses" of $\mathfrak{S}$; the degree of a rational function, by the "rank" of an element of $\mathfrak{S}$; and the product of two differential expressions, by the "product" of two "parentheses," the abstract "product" being subjected to postulates corresponding to the properties (*a*) and (*b*).

## 2. The abstract theory $\mathfrak{T}_{-1}$

Let $\mathfrak{S}$ denote any given *aggregate* whatsoever.† We shall deal with finite, ordered subaggregates $E = (e_0, e_1, e_2, \cdots, e_m)$ of $\mathfrak{S}$, $e_\nu$ $\{\nu = 0, 1, 2, \cdots, m\}$ being an element of $\mathfrak{S}$. Such a finite, ordered subaggregate $E$ we shall call a "*parenthesis* of $\mathfrak{S}$." We assume that the "*product*"

$$A \cdot B = (a_0, a_1, a_2, \cdots, a_r) \cdot (b_0, b_1, b_2, \cdots, b_s)$$

of any two parentheses $A$ and $B$ of $\mathfrak{S}$ is equal to a parenthesis

$$C = (c_0, c_1, c_2, \cdots, c_n)$$

of $\mathfrak{S}$ for which $n = r + s$. We assume furthermore that with every element $e$ of $\mathfrak{S}$ there is associated a single number $\eta$ called the "*rank*" of $e$, $\eta$ being a real number or $-\infty$ (never $+\infty$),‡ such that when $A \cdot B = C$

* The reason for the notation "$\mathfrak{T}_{-1}$" lies, as indicated in the introduction, in the fact that this abstract theory is later (§ 8) exhibited as a special case of a more general theory $\mathfrak{T}_\epsilon$, the latter reducing to the former if $\epsilon = -1$.

† In regard to the use of the unrestricted aggregate (= general range) and several other features of this paper, such as the postulational unification of analogous theories, compare E. H. Moore, *Introduction to a Form of General Analysis* (1910), preface.

‡ In all the concrete situations 1–10 incl., with the exception of 4*b*, 7*b*, and 9*b*, $\eta$ happens to be always an *integer*. The proof of the "Theorem on the Product" is, however, not rendered more difficult if $\eta$ is merely restricted to being real.

the following properties hold for the "ranks" $\alpha_\lambda, \beta_\mu, \gamma_\nu$ of $a_\lambda, b_\mu, c_\nu$, $\{\nu = 0, 1, 2, \cdots, n\}$:*

(a) $\gamma_\nu \leqq \max_{\lambda+\mu+\sigma=\nu} (\alpha_\lambda + \beta_\mu - \sigma)$, $\{0 \leqq \lambda \leqq r, 0 \leqq \mu \leqq s, 0 \leqq \sigma\}$.

(b) $\gamma_\nu = \alpha_{\bar\lambda} + \beta_{\bar\mu}$ $\{\bar\lambda + \bar\mu = \nu\}$ if $\alpha_{\bar\lambda} + \beta_{\bar\mu} > \alpha_\lambda + \beta_\mu - \sigma$ $\{\lambda + \mu + \sigma = \nu\}$ for every $[\lambda, \mu, \sigma] \neq [\bar\lambda, \bar\mu, 0]$, $\{0 \leqq \lambda \leqq r, 0 \leqq \mu \leqq s, 0 \leqq \sigma\}$.

The "*order*" of the parenthesis $E \equiv (e_0, e_1, \cdots, e_m)$ is defined as $n - k$, where $k$ is such that $\eta_0 = -\infty, \eta_1 = -\infty, \cdots, \eta_{k-1} = -\infty$, and $\eta_k \neq -\infty$. If every $\eta_\nu$ $\{\nu = 0, 1, \cdots, m\}$ is $-\infty$, the order of $E$ is defined to be 0.

"Reducibility," "irreducibility," "factor," etc., are defined for $\mathfrak{T}_{-1}$ (and later for $\mathfrak{T}_\epsilon$) in expected fashion. Thus $C$ is said to be "*reducible*" if two parentheses $A$ and $B$ of $\mathfrak{S}$ exist, such that $r \geqq 1$, $s \geqq 1$ and $C = A \cdot B$.

### 3. Instances of the abstract theory $\mathfrak{T}_{-1}$

We shall now describe various important situations that fall under the abstract theory $\mathfrak{T}_{-1}$. For this purpose, we shall define in each case the *aggregate* $\mathfrak{S}$, the *parentheses* of $\mathfrak{S}$, the *product* of two parentheses and the *rank* of every element of $\mathfrak{S}$. The verification of the properties (*a*) and (*b*) of § 2 for the instances given below is left to the reader.

*Situation* 1. $\mathfrak{S}$ consists of the set of rational numbers. We understand by the parenthesis $(e_0, e_1, e_2, \cdots, e_m)$ of $\mathfrak{S}$ the polynomial

$$e_0 y^m + e_1 y^{m-1} + \cdots + e_m$$

in the letter $y$. The product of two parentheses

$$(a_0, a_1, \cdots, a_r) \cdot (b_0, b_1, \cdots, b_s)$$

is defined, as usual, to be equal to the parenthesis $(c_0, c_1, \cdots, c_n)$, where $n = r + s$ and

$$c_0 = a_0 b_0, \qquad c_1 = a_0 b_1 + a_1 b_0, \qquad \cdots, \qquad c_n = a_r b_s.$$

The rank of an element $e = e'/e''$ (where $e'$ and $e''$ are integers) is defined with reference to a fixed prime $p$. First let $e \neq 0$; let $e'$ be divisible by $p^{d'}$ but not by $p^{d'+1}$; $e''$, by $p^{d''}$ but not by $p^{d''+1}$. We then define the rank of $e$ by the equation $\eta = d'' - d'$. Moreover, we naturally enough define the rank of 0 to be $-\infty$.

*Situation* 2. $\mathfrak{S}$ consists of the class of Hensel's $p$-adic numbers. The parentheses of $\mathfrak{S}$ and the products of parentheses are defined as in 1. The

* Because of the nature of our assumptions for $\mathfrak{T}_{-1}$, which involve essentially only the ranks of the elements of $\mathfrak{S}$, it is possible to build up just as general a theory as $\mathfrak{T}_{-1}$ by dealing exclusively with parentheses whose elements are ranks and hence always real numbers or $-\infty$. We have here an example of an apparently very general theory (based on an unrestricted aggregate and involving few assumptions) that is equivalent to an apparently very special one (about parentheses of real numbers).

rank of the $p$-adic number $e$ is the negative of Hensel's "order" of $e$ with respect to $p$.*

*Situation* 3. $\mathfrak{S}$ consists of the class of rational fractions

$$e = e'(x_1, \cdots, x_k)/e''(x_1, \cdots, x_k)$$

in $k$ letters, $x_1, x_2, \cdots, x_k$, $e'$ and $e''$ being polynomials with arbitrary complex numerical coefficients.† The parentheses and the products of parentheses are defined as in 1. The rank of an element $e(\neq 0)$ is defined by the equation $\eta = d' - d''$, where $d'$ and $d''$ represent respectively the degrees (in the usual sense) of $e'$ and $e''$ in the $k$ letters $x_1, x_2, \cdots, x_k$. The rank of 0 is defined to be $-\infty$.

*Situation* 4*a*. $\mathfrak{S}$ consists of the collectivity of elements $e = \sum_{\nu=0}^{\nu=\infty} v_\nu x^{\eta-\nu}$, where the $v$'s are arbitrary complex numbers, $\eta$ is an integer and also $v_0 \neq 0$ unless $e = 0$; that is, $e$ is a Laurent series having only a finite number of terms with positive exponents.‡ The parentheses and the products are defined as in 1. The rank of $e(\neq 0)$ is defined to be the exponent $\eta$ of $x$ in the first non-zero term of the development of $e$. The rank of 0 is $-\infty$.

*Situation* 4*b*. $\mathfrak{S}$ consists of the collectivity of series $\sum_{\nu=0}^{\nu=\infty} v_\nu x^{(\theta/\delta)-(\nu/\delta)}$ where $\theta$ and $\delta$ $(>1)$ are integers; the elements of $\mathfrak{S}$ are thus descending series in integral powers of $x^{1/\delta}$. Otherwise, as in 4*a*.

*Situation* 5. $\mathfrak{S}$ consists of the collectivity of fractions $e = e'/e''$, where $e'$ and $e''$ are power series in $k$ letters $x_1, x_2, \cdots, x_k$ with arbitrary complex numerical coefficients:

$$e' = \sum_{\lambda_1, \cdots, \lambda_k=0}^{\infty} v'_{\lambda_1, \cdots, \lambda_k} x_1^{\lambda_1} \cdots x_k^{\lambda_k},$$

$$e'' = \sum_{\lambda_1, \cdots, \lambda_k=0}^{\infty} v''_{\lambda_1, \cdots, \lambda_k} x_1^{\lambda_1} \cdots x_k^{\lambda_k}.$$

The parentheses and products are defined as in 1. The rank of $e$ $(\neq 0)$ is defined as $d'' - d'$, where $d'$ and $d''$ represent the lowest degrees (in the usual sense) in the $k$ letters $x_1, \cdots, x_k$ of a non-zero term of $e'$ and $e''$ respectively. The rank of 0 is $-\infty$.

*Situation* 6.§ $\mathfrak{S}$ consists of all rational fractions $e(x) = e'(x)/e''(x)$, $e'(x)$ and $e''(x)$ being polynomials with arbitrary complex numerical coef-

* See Hensel, Journal für Mathematik, vol. 127 (1904), pp. 51–84, or *Zahlentheorie* (1913), chapters 3 and 6.

† More generally, the numerical coefficients may belong to any abstract system $(K, +, \times)$, where $K$ is a class such that if $a$ and $b$ are elements of $K$ both $a + b$ and $a \times b$ are elements of $K$. This remark applies just as well to other situations below.

‡ The question of convergence does not enter here because the formal character of the series is sufficient for our purposes. More generally, we may have similar series in two or more letters, the requisite change in the definition of rank being evident.

§ Situation 6 is the illustrative example of § 1.

ficients. The parenthesis $(e_0, e_1, \cdots, e_m)$ is the linear homogeneous differential expression

$$e_0(x)\frac{d^m y}{dx^m} + e_1(x)\frac{d^{m-1} y}{dx^{m-1}} + \cdots + e_m(x)\, y.$$

The rank of $e$ is the degree of $e$ (in the usual sense), and the product of two parentheses is the *symbolic* product of the two corresponding differential expressions.

*Situation* 7*a*. $\mathfrak{S}$ and rank are defined as in 4*a*; parentheses and products, as in 6.

*Situation* 7*b*. $\mathfrak{S}$ and rank are defined as in 4*b*; parentheses and products, as in 6.

*Situation* 8. $\mathfrak{S}$ and rank are defined as in 6. The parenthesis

$$(e_0, e_1, \cdots, e_m)$$

of $\mathfrak{S}$ is the linear homogeneous difference expression

$$e_0\, y_{x+m} + e_1\, y_{x+m-1} + \cdots + e_m\, y_x.$$

The product $A \cdot B$ of two parentheses is the ordinary *symbolic* product of the corresponding difference expressions:

$$A \cdot B = (a_0, a_1, \cdots, a_r) \cdot (b_0, b_1, \cdots, b_s) = (c_0, c_1, \cdots, c_n),$$

where $n = r + s$ and

$$c_\nu = \sum_{\rho=0}^{\nu} a_\rho(x)\, b_{\nu-\rho}(x + r - \rho), \qquad \{\nu = 0, 1, 2, \cdots, n\}.$$

*Situation* 9*a*. $\mathfrak{S}$ and rank are defined as in 4*a*; parentheses and products, as in 8.

*Situation* 9*b*. $\mathfrak{S}$ and rank are defined as in 4*b*; parentheses and products, as in 8.

## 4.* The characteristic set $S_C$ belonging to $C$

From now on (till § 8), our discussion is based on the assumptions for the abstract theory $\mathfrak{T}_{-1}$ of § 2. *In the present section, however (and also in* § 5*), we make use merely of the fact that every element of* $\mathfrak{S}$ *has a rank* (which is a real number or $-\infty$).†

Let $C = (c_0, c_1, \cdots, c_n)$ be a given parenthesis of $\mathfrak{S}$, $\gamma_\nu$ being the rank of $c_\nu$ $\{\nu = 0, 1, \cdots, n\}$. *We assume throughout in what follows that* $\gamma_0$ *is finite*

* At this point, the reader may pass directly to § 8 for the definition of $\mathfrak{T}_\epsilon$; he will thereby be enabled (on second reading, after §§ 4–7 incl.) to verify more easily the statements of § 8.

† On the other hand, if we have no regard to (*a*) and (*b*) of § 2, the definition of "characteristic set" here given appears highly artificial.

(i. e., $\neq -\infty$). Let, moreover, $n \geqq 1$ (a condition to be subsequently dropped). With every such parenthesis $C$ we shall associate in unique fashion an ordered set $S_C$ of number pairs:

$$S_C = \{(m_1, t_1), (m_2, t_2), \cdots, (m_h, t_h)\}.$$

This set, which we call the "characteristic set belonging to $C$," we now define.* For short, we designate $\sum_{\nu=1}^{\nu=\zeta} t_\nu$ by $s_\zeta$ $\{\zeta = 1, 2, \cdots, h\}$; $s_0 = 0$. We set $m_1$ equal to $\max_{\nu>0}(\gamma_\nu - \gamma_0/\nu)$ except when this maximum is $< -1$; in the latter case, we set $m_1 = -1$. If $m_1 > -1$, we set $t_1$ equal to the last $\nu$ $(\nu = 1, 2, \cdots, n)$ for which $(\gamma_\nu - \gamma_0)/\nu = m_1$; and if $m_1 = -1$, we set $t_1 = n$.† In general, we define $m_{\zeta+1}$ and $s_{\zeta+1}$ (and hence $t_{\zeta+1}$) for $\zeta = 0, 1, \cdots, h-1$ as follows:

$$m_{\zeta+1} = \max_{\nu>s_\zeta}\left(\frac{\gamma_\nu - \gamma_{s_\zeta}}{\nu - s_\zeta}, -1\right).$$

$$\begin{cases} s_{\zeta+1} = \textit{last } \nu > s_\zeta \textit{ such that } \dfrac{\gamma_\nu - \gamma_{s_\nu}}{\nu - s_\zeta} = m_{\zeta+1}, \textit{ if } m_{\zeta+1} > -1; \textit{ and} \\ s_{\zeta+1} = n, \textit{ if } m_{\zeta+1} = -1. \end{cases}$$

The following relations are direct consequences of the definition of $S_C$:

(a) $s_h = \sum_{\nu=1}^{h} t_\nu = n; \qquad t_\nu \geqq 1 \qquad \{\nu = 1, 2, \cdots, h\}.$

(b) $m_1 > m_2 > \cdots > m_h \geqq -1.$

(1) (c) $$\frac{\gamma_\nu - \gamma_{s_\zeta}}{\nu - s_\zeta}\{\nu > s_\zeta\} \leqq m_{\zeta+1} > \frac{\gamma_\nu - \gamma_{s_\zeta}}{\nu - s_\zeta}\{\nu > s_{\zeta+1}\}, \quad \{\zeta = 0, 1, 2, \cdots, h-1\}.$$

(d) $\gamma_{s_\zeta} - \gamma_{s_{\zeta-1}} = m_\zeta t_\zeta$ *if* $m_\zeta > -1$; *hence*

$$\gamma_{s_\zeta} - \gamma_0 = \sum_{\nu=1}^{\zeta} m_\nu t_\nu \textit{ if } m_\zeta > -1.$$

As in the theory of well-ordered sets, we call the set

* This "characteristic set" is the natural generalization of the "*polygon of Newton*" or of the "*Puiseux numbers*" for our abstract situation $\mathfrak{T}_{-1}$. The "generalized polygon of Newton" (for $\mathfrak{T}_{-1}$) may be essentially obtained by starting at the origin of a rectangular system of axes, and drawing a broken line such that the slope of the $\zeta$th segment is $m_\zeta$ and its projection on the $X$-axis, $t_\zeta$. This broken line is, according to 1*b* below, concave downward. The author might have made the entire treatment (including the proof of the "Theorem on the Product") depend on the properties of this broken line; he has chosen, however, the equivalent analytical procedure. Cf. Hensel and Landsberg, *Theorie der Algebraischen Funktionen einer Variabeln* (1902), vierte Vorles., pp. 39–47 and Bauer, Journal für Mathematik, vol. 132, pp. 21–32.

† In the theory $\mathfrak{T}_{-1}$, the number $-1$ plays an exceptional rôle; in the theory $\mathfrak{T}_\epsilon$ (see § 8) the number $\epsilon$ plays the corresponding rôle.

$$S_{C_\zeta} = \{(m_1, t_1), \cdots, (m_\zeta, t_\zeta)\}$$

for $\zeta = 1, 2, \cdots, h$ a "*segment*" of $S_C$. The set $S_{C_\zeta}$ is evidently the characteristic set of the parenthesis $(c_0, c_1, \cdots, c_{s_\zeta})$ of $\mathfrak{S}$. It is convenient in the sequel to *adjoin the null-segment* $0 = S_{C_0}$ *of* $S_C$ *and regard it as the characteristic set of every parenthesis* $(c_0)$ of order 0. As indicated above, we thus extend the definition of characteristic set to the case $n = 0$. Every characteristic set has one and the same null-segment 0.

LEMMA* I. *For every* $(\nu, \zeta)$, $\{\nu = 0, 1, 2, \cdots, n;\ \zeta = 0, 1, 2, \cdots, h\}$ *such that* $m_\zeta > -1$, *we have* $\gamma_\nu - \gamma_{s_\zeta} \leqq (\nu - s_\zeta)\, m_\zeta$ *and also* $\leqq (\nu - s_\zeta)\, m_{\zeta+1}$.

*Proof.* If $\nu = s_\zeta$, the lemma is trivial. If $\nu > s_\zeta$ (whence $m_{\zeta+1}$ exists), the lemma follows from 1*c* and 1*b*. If $\nu < s_\zeta$ (whence $m_\zeta$ exists), let $\omega$ be the uniquely determined integer satisfying the relations $s_{\omega-1} \leqq \nu < s_\omega$. Then $1 \leqq \omega \leqq \zeta$. According to the part of the lemma already established,

$$\gamma_\nu - \gamma_{s_{\omega-1}} \leqq (\nu - s_{\omega-1})\, m_\omega .$$

Moreover,

$$\gamma_{s_\zeta} - \gamma_{s_{\omega-1}} = \sum_{\nu=\omega}^{\zeta} m_\nu t_\nu , \tag{1d}$$

Hence

$$\begin{aligned} \gamma_\nu - \gamma_{s_\zeta} &\leqq (\nu - s_\omega)\, m_\omega - \sum_{\nu=\omega+1}^{\zeta} m_\nu t_\nu \\ &\leqq \left(\nu - s_\omega - \sum_{\nu=\omega+1}^{\zeta} t_\nu\right) m_\zeta \\ &= (\nu - s_\zeta)\, m_\zeta . \end{aligned} \tag{1b}$$

*Remark.* If $\nu > s_\zeta$, $\gamma_\nu - \gamma_{s_\zeta} < (\nu - s_\zeta)\, m_\zeta$ (and, in fact, $< (\nu - s_\zeta)\, m$ for every $m > m_{\zeta+1}$); and if $\nu < s_\zeta$, $\gamma_\nu - \gamma_{s_\zeta} < (\nu - s_\zeta)\, m_{\zeta+1}$ (and, in fact, $< (\nu - s_\zeta)\, m$ for every $m < m_\zeta$).

## 5. THE SUM OF TWO CHARACTERISTIC SETS

Let $A = (a_0, a_1, \cdots, a_r)$ and $B = (b_0, b_1, \cdots, b_s)$ be two parentheses of $\mathfrak{S}$, $\alpha_\lambda$ and $\beta_\mu$ being the rank of $a_\lambda$ and $b_\mu$ respectively; furthermore, let $\alpha_0 \neq -\infty$ and $\beta_0 \neq -\infty$. Then, according to the previous section, the characteristic sets

$$S_A = \{(m_1', t_1'), \cdots, (m_f', t_f')\}$$

* *The occurrence of* $m_\zeta$ $(m_{\zeta+1})$ *in the lemma is not to be interpreted as implying that* $m_\zeta$ $(m_{\zeta+1})$ *exists*; for it is our purpose, as shown in the formulation of the lemma, to include also the cases $\zeta = 0$, when $m_0$ does not exist, and $\zeta = h$, when $m_{\zeta+1}$ does not exist. If $m_\zeta$ $(m_{\zeta+1})$ does not exist, every relation rendered meaningless because of this non-existence is *regarded as being fulfilled vacuously*. Thus, if $\zeta = 0$, $m_0$ is non-existent and $m_0 > -1$ is regarded as true. On the other hand, every relation that has meaning requires proof. *In the light of these remarks, similar situations below—and a few above—are to be interpreted.*

and

$$S_B = \{(m_1'', t_1''), \cdots, (m_g'', t_g'')\}$$

exist. The sum $S_A + S_B$ is defined (for $r \geqq 1$ and $s \geqq 1$) as a new set of number pairs

$$S_C = \{(m_1, t_1), \cdots, (m_h, t_h)\},$$

the $m$'s and the $t$'s being defined in the following manner:* The sequence $\{m_1, m_2, \cdots, m_h\}$ is obtained by arranging the (single) set of numbers $m_\nu'$ $\{\nu = 1, 2, \cdots, f\}$ and $m_\nu''$ $\{\nu = 1, 2, \cdots, g\}$ in a decreasing sequence according to magnitude and discarding duplicates. With each $m$ associate the number $t = t_\mu'$ or $t = t_\nu''$ that corresponds to it in either $S_A$ or $S_B$ if there is just one such correspondent; if there are two, $t_\mu'$ in $S_A$ and $t_\nu''$ in $S_B$, associate with $m$ the number $t = t_\mu' + t_\nu''$. Furthermore, $S_A + S_B$ is defined to be $S_B$ or $S_A$ according as $S_A = 0$ (which is the case when $r = 0$) or $S_B = 0$ (which is the case when $s = 0$).

From the definition of $S_A + S_B$ it follows that to every segment

$$S_{C_\zeta} = \{(m_1, t_1), \cdots, (m_\zeta, t_\zeta)\}$$

of $S_C$, where $\zeta = 0, 1, 2, \cdots, h$, there correspond uniquely two segments

$$S_{A_\xi} = \{(m_1', t_1'), \cdots, (m_\xi', t_\xi')\}$$

of $S_A$ and

$$S_{B_\eta} = \{(m_1'', t_1''), \cdots, (m_\eta'', t_\eta'')\}$$

of $S_B$ such that $S_{A_\xi} + S_{B_\eta} = S_{C_\zeta}$. Two segments $S_{A_\xi}$ and $S_{B_\eta}$ thus determined are said to be "*conjugate.*" There are thus in all $h + 1$ pairs of conjugate segments.

*Two conjugate segments $S_{A_\xi}$ and $S_{B_\eta}$ are characterized by the property that $m_\xi' > m_{\eta+1}''$ and $m_\eta'' > m_{\xi+1}'$.* (The null-segments and the segments $S_{A_f} = S_A$, $S_{B_g} = S_B$ share this property on account of our agreement regarding vacuous validity.)

It is evident that the $h + 1$ pairs of conjugate segments may be so ordered, —and in only one way— that $\xi \leqq \xi'$ and $\eta \leqq \eta'$ (and furthermore, of course, either $\xi < \xi'$ or $\eta < \eta'$), whenever the pair $(S_{A_\xi}, S_{B_\eta})$ precedes the pair $(S_{A_{\xi'}}, S_{B_{\eta'}})$. Then $(S_{A_0}, S_{B_0})$ is the first pair, and $(S_{A_f}, S_{B_g}) \equiv (S_A, S_B)$, the last.

We introduce here, with regard to conjugate segments, a lemma that will simplify the proof of the "Theorem on the Product." This lemma gives us desirable information regarding the comparative values of the expressions $\alpha_\lambda + \beta_\mu - \sigma$ and $\alpha_{s_\xi'}' + \beta_{s_\eta''}''$ (see § 2). Here $s_\xi'$, $s_\eta''$ stand, of course, for $\sum_{\nu=1}^{\nu=\xi} t_\nu'$, $\sum_{\nu=1}^{\nu=\eta} t_\nu''$.

* The designation $S_C$ is not meant to imply here the existence of a parenthesis $C$ having $S_C$ as characteristic set.

LEMMA II. *Let* $A(\alpha_0 \neq -\infty)$ *and* $B(\beta_0 \neq -\infty)$ *be two parentheses of* $\mathfrak{S}$; $S_{A_\xi}$, $S_{B_\eta}$, *two conjugate segments of* $S_A$, $S_B$; *and* $\lambda$, $\mu$, $\sigma$ *integers such that* $0 \leqq \lambda \leqq r$, $0 \leqq \mu \leqq s$ *and* $0 \leqq \sigma$. *Set*

$$\delta = \lambda + \mu + \sigma - (s'_\xi + s''_\eta)$$

*and*

$$\Delta = \alpha_\lambda + \beta_\mu - \sigma - (\alpha_{s'_\xi} + \beta_{s''_\eta}).$$

*Then*

(*a*) *if* $\delta > 0$ *and* $m = \max(m'_{\xi+1}, m''_{\eta+1})$, $\Delta \leqq \delta m$;

(*b*) *if* $\delta = 0$, $m'_\xi > -1$, $m''_\eta > -1$ *and* $[\lambda, \mu, \sigma] \not\equiv [s'_\xi, s''_\eta, 0]$, $\Delta < 0$.

*Proof of* (*a*). On account of our agreement regarding vacuous validity, there is nothing to prove if both $m'_{\xi+1}$ and $m''_{\eta+1}$ are non-existent.* On account of symmetry, we may then assume without losing generality that $m'_{\xi+1}$ exists and $\geqq m''_{\eta+1}$. From Lemma I, we see that $\alpha_\lambda - \alpha_{s'_\xi} \leqq (\lambda - s'_\xi) m'_{\xi+1}$; furthermore, since $S_{A_\xi}$ and $S_{B_\eta}$ are conjugate, $m''_\eta > m'_{\xi+1} \geqq m''_{\eta+1}$ and therefore $\beta_\mu - \beta_{s''_\eta} \leqq (\mu - s''_\eta) m'_{\xi+1}$; finally, since $-1 \leqq m'_{\xi+1}$, $-\sigma \leqq \sigma m'_{\xi+1}$. Accordingly,

$$\Delta \leqq (\lambda - s'_\xi + \mu - s''_\eta + \sigma) m'_{\xi+1} = \delta m.$$

*Proof of* (*b*). We suppose that $m'_\xi$ and $m''_\eta$ both exist, the requisite modification of the argument being evident when this supposition is not fulfilled. Without loss of generality, we may assume, on account of symmetry, that $m'_\xi \geqq m''_\eta$. We divide our discussion into two parts according as $\sigma > 0$ or $\sigma = 0$.

If $\sigma > 0$, $-\sigma < \sigma m''_\eta$, since $-1 < m''_\eta$ by hypothesis. Furthermore, since $S_{A_\xi}$ and $S_{B_\eta}$ are conjugate segments, $m'_\xi \geqq m''_\eta > m'_{\xi+1}$; hence (Lemma I)

$$\alpha_\lambda - \alpha_{s'_\xi} \leqq (\lambda - s'_\xi) m''_\eta.$$

Since, finally, $\beta_\mu - \beta_{s''_\eta} \leqq (\mu - s''_\eta) m''_\eta$, we have

$$\Delta < (\lambda - s'_\xi + \mu - s''_\eta + \sigma) m''_\eta = 0.$$

If $\sigma = 0$, it follows from $[\lambda, \mu, \sigma] \not\equiv [s'_\xi, s''_\eta, 0]$ and $\delta = 0$ that either (1) $\lambda - s'_\xi > 0$ and $\mu - s''_\eta < 0$ or (2) $\lambda - s'_\xi < 0$ and $\mu - s''_\eta > 0$. In the first case, $m'_{\xi+1}$ exists and (Lemma I, Remark)

$$\alpha_\lambda - \alpha_{s'_\xi} \leqq (\lambda - s'_\xi) m'_{\xi+1} < (\lambda - s'_\xi) m''_\eta.$$

Since, moreover, $\beta_\mu - \beta_{s''_\eta} \leqq (\mu - s''_\eta) m''_\eta$, it follows that

* As indicated in a previous footnote (§ 4), it is not presupposed in Lemma II that $m'_\xi$, $m'_{\xi+1}$, $m''_\eta$, $m''_{\eta+1}$ exist. It is our purpose to include in the application of the lemma the cases $\xi = 0$, $\eta = 0$, $\xi = f$, $\eta = g$. If $m'_{\xi+1}$ exists but $m''_{\eta+1}$ does not, it is to be understood that $\max(m'_{\xi+1}, m''_{\eta+1}) = m'_{\xi+1}$; similarly, if $m''_{\eta+1}$ exists and $m'_{\xi+1}$ does not. Since the relation $\Delta < 0$ has meaning in (*b*), even when $m'_\xi$ or $m''_\eta$ or both are non-existent, it requires proof also in the latter cases.

$$\Delta < (\lambda - s'_\xi + \mu - s''_\eta)\, m''_\eta = 0.$$

In the second case, $\beta_\mu - \beta_{s''_\eta} < (\mu - s''_\eta)\, m''_\eta$ (Lemma I, Remark). Since, moreover,

$$\alpha_\lambda - \alpha_{s'_\xi} \leqq (\lambda - s'_\xi)\, m'_\xi \leqq (\lambda - s'_\xi)\, m''_\eta$$

(on account of $\lambda - s'_\xi < 0$ and $m''_\eta \leqq m'_\xi$), $\Delta < 0$.

## 6. The "theorem on the product"*

THEOREM. *If $A(\alpha_0 \neq -\infty)$ and $B(\beta_0 \neq -\infty)$ are two parentheses of $\mathfrak{S}$ and $C = A \cdot B$, then $S_C = S_A + S_B$.*

*Proof.* $C$ is a parenthesis of $\mathfrak{S}$ by virtue of § 2. Moreover,

$$\gamma_0 = \alpha_0 + \beta_0. \tag{1}$$

For, if in (*b*) § 2 we set $\nu = 0$, it follows, on account of $0 \leqq \lambda, 0 \leqq \mu, 0 \leqq \sigma$, that $\lambda = 0$, $\mu = 0$, $\sigma = 0$ for every $[\lambda, \mu, \sigma]_{\lambda+\mu+\sigma=\nu}$; hence there is no $[\lambda, \mu, \sigma]_{\lambda+\mu+\sigma=\nu} \neq [\bar{\lambda}, \bar{\mu}, 0] \equiv [0, 0, 0]$. The hypothesis of (*b*) § 2 is therefore fulfilled vacuously, and in consequence (1) holds. From (1) it follows that $\gamma_0$ is finite and accordingly that $S_C$ exists.

The proof that $S_C = S_A + S_B$ is effected by means of mathematical induction. We thus assume that it has already been established for the conjugate segments $S_{A_\xi}$, $S_{B_\eta}$ of $S_A$, $S_B$ that a segment $S_{C_\zeta}$ of $S_C$ exists such that

$$S_{C_\zeta} = S_{A_\xi} + S_{B_\eta}. \tag{2}$$

Since $S_{C_0} = S_{A_0} + S_{B_0}$ (each term in this equation being the null-set), the theorem will be completely proved, if, on the basis of the assumption (2), we show that $m_{\zeta+1}$ and $t_{\zeta+1}$ have values in accordance with the implications of $S_C = S_A + S_B$. (Compare the ordering of the $h+1$ pairs of conjugate segments, mentioned in § 5.) This is what we shall do.

From (2) it follows, in accordance with the definition of the sum of two characteristic sets, that

$$s_\zeta = s'_\xi + s''_\eta. \tag{3}$$

There is nothing to prove if neither $m'_{\xi+1}$ nor $m''_{\eta+1}$ exists; for then, according to (3) and § 2,

$$s_\zeta = s'_f + s''_g = r + s = n,$$

and therefore $m_{\zeta+1}$ does not exist. Hence we may also assume, without loss of generality, that $m'_\xi > -1$ and $m''_\eta > -1$. For suppose $m'_\xi = -1$ (a similar argument will hold if $m''_\eta = -1$); then $m'_{\xi+1}$ does not exist. Moreover, on account of the conjugacy of $S_{A_\xi}$ and $S_{B_\eta}$,

* This designation was suggested by Professor E. W. Davis.

$$-1 = m'_\xi > m''_{\eta+1} \geqq -1.$$

Hence $m''_{\eta+1}$ does not exist (or else we may conclude that $-1 > -1$).* From $m'_\xi > -1$ and $m''_\eta > -1$, it follows, however, by virtue of 1*d*, § 4, and the relations (1) and (2) of the present section, that

$$\gamma_{s_\zeta} = \alpha_{s'_\xi} + \beta_{s''_\eta}. \tag{4}$$

On account of symmetry, we may furthermore assume, without loss of generality, that $m'_{\xi+1}$ exists and $\geqq m''_{\eta+1}$. We now divide our discussion (in natural manner) into the following three parts:

(*a*) $m'_{\xi+1} = -1$. According to (*a*) Lemma II, if

$$\nu = \lambda + \mu + \sigma > s'_\xi + s''_\eta = s_\zeta,$$

we have, since $\delta = \nu - s_\zeta$, $\alpha_{s'_\xi} + \beta_{s''_\eta} = \gamma_{s_\zeta}$, and $m = m'_{\xi+1} = -1$:

$$\alpha_\lambda + \beta_\mu - \sigma \leqq \gamma_{s_\zeta} - (\nu - s_\zeta).$$

Hence, by virtue of (*a*) § 2, $\gamma_\nu \leqq \gamma_{s_\zeta} - (\nu - s_\zeta)$ for every $\nu > s_\zeta$; i. e.,

$$\frac{\gamma_\nu - \gamma_{s_\zeta}}{\nu - s_\zeta}\{\nu > s_\zeta\} \leqq -1.$$

According to the definition of " characteristic set," this shows that $m_{\zeta+1} = -1$ and $s_{\zeta+1} = n = r + s$; that is, by (3), $t_{\zeta+1} = t'_{\xi+1}$ or $t'_{\xi+1} + t''_{\eta+1}$ according as $m''_{\eta+1}$ is non-existent or existent (the value of $m''_{\eta+1}$ being $-1$ in the latter case).

(*b*) $m'_{\xi+1} > -1$ and $> m''_{\eta+1}$ ($m''_{\eta+1}$ being existent or non-existent). By means of (*a*) Lemma II, we show, as in the case $m'_{\xi+1} = -1$, that

$$\frac{\gamma_\nu - \gamma_{s_\zeta}}{\nu - s_\zeta}\{\nu > s_\zeta\} \leqq m'_{\xi+1}.$$

Moreover, the segments $S_{A_{\xi+1}}$ and $S_{B_\eta}$ are conjugate; for, in addition to $m'_{\xi+1} > m''_{\eta+1}$, we may conclude, on account of the assumed conjugacy of $S_{A_\xi}$ and $S_{B_\eta}$ and 1*b*, § 4, that $m''_\eta > m'_{\xi+1} > m'_{\xi+2}$ (the latter relation holding vacuously when $m'_{\xi+2}$ does not exist). It therefore follows from (*b*) Lemma II and (*b*) § 2 that

$$\gamma_\nu\{\nu = s'_{\xi+1} + s''_\eta\} = \alpha_{s'_{\xi+1}} + \beta_{s''_\eta} = \gamma_{s_\zeta} + t'_{\xi+1} m'_{\xi+1} = \gamma_{s_\zeta} + (\nu - s_\zeta) m'_{\xi+1}$$

(1*d*, § 4 and (4) of present section),

whence

$$\frac{\gamma_\nu - \gamma_{s_\zeta}}{\nu - s_\zeta}\{\nu = s'_{\xi+1} + s''_\eta\} = m'_{\xi+1}.$$

* If $m'_\xi$ does not exist (as when $\xi = 0$), the contradiction $-1 > -1$ cannot be obtained even if $m''_{\eta+1}$ exists; but in that case, $m'_\xi > -1$ holds vacuously.

In view of the definition of "characteristic set," we may now conclude that $m_{\zeta+1} = m'_{\xi+1}$. If neither $m'_{\xi+2}$ nor $m''_{\eta+1}$ exists, $t_{\zeta+1}$ is evidently $t'_{\xi+1}$ and there is nothing more to prove. We may therefore assume without losing generality that either $m'_{\xi+2}$ or $m''_{\eta+1}$ exists; hence $m = \max(m'_{\xi+2}, m''_{\eta+1})$ exists. According to (*a*) Lemma II, we then see, in virtue of the conjugacy of $S_{A_{\xi+1}}$ and $S_{B_\eta}$, that for $\nu = \lambda + \mu + \sigma > s'_{\xi+1} + s''_\eta$, we have

$$\alpha_\lambda + \beta_\mu - \sigma \leqq \alpha_{s'_{\xi+1}} + \beta_{s''_\eta} + (\nu - s'_{\xi+1} - s''_\eta)\, m.$$

But

$$\alpha_{s'_{\xi+1}} + \beta_{s''_\eta} = \alpha_{s'_\xi} + \beta_{s''_\eta} + t'_{\xi+1}\, m'_{\xi+1} = \gamma_{s_\zeta} + t'_{\xi+1}\, m'_{\xi+1}$$

(1*d*, § 4, since $m'_{\xi+1} > -1$),

by (4); furthermore,

$$(\nu - s'_{\xi+1} - s''_\eta)\, m < (\nu - s'_{\xi+1} - s''_\eta)\, m'_{\xi+1},$$

since $\nu - s'_{\xi+1} - s''_\eta$ is positive and both $m'_{\xi+2}$ and $m''_{\eta+1}$ are $< m'_{\xi+1}$. Therefore

$$\alpha_\lambda + \beta_\mu - \sigma < \gamma_{s_\zeta} + (\nu - s'_\xi - s''_\eta)\, m'_{\xi+1} = \gamma_{s_\zeta} + (\nu - s_\zeta)\, m'_{\xi+1},$$

by (3). Consequently, in accordance with (*a*), § (2),

$$\gamma_\nu < \gamma_{s_\zeta} + (\nu - s_\zeta)\, m'_{\xi+1},$$

whence

$$\frac{\gamma_\nu - \gamma_{s_\zeta}}{\nu - s_\zeta} \{\nu > s'_{\xi+1} + s''_\eta\} < m'_{\xi+1}.$$

We now see that $s_{\zeta+1} = s'_{\xi+1} + s''_\eta$ and hence $t_{\zeta+1} = t'_{\xi+1}$.

(*c*) $m'_{\xi+1} = m''_{\eta+1} > -1$. As in the previous case,

$$\frac{\gamma_\nu - \gamma_{s_\zeta}}{\nu - s_\zeta} \{\nu > s_\zeta = s'_\xi + s''_\eta\} \leqq m'_{\xi+1};$$

moreover, by means of (*b*) Lemma II for the conjugate segments $S_{A_{\xi+1}}$ and $S_{B_{\eta+1}}$ and (*b*) § 2, it follows that

$$\frac{\gamma_\nu - \gamma_{s_\zeta}}{\nu - s_\zeta} \{\nu = s'_{\xi+1} + s''_{\eta+1}\} = m'_{\xi+1};$$

and finally, by means of (*a*) Lemma II for these segments and (*a*) § 2 (as in the previous case), that

$$\frac{\gamma_\nu - \gamma_{s_\zeta}}{\nu - s_\zeta} \{\nu > s'_{\xi+1} + s''_{\eta+1}\} < m'_{\xi+1}.$$

From these relations we conclude that $m_{\zeta+1} = m'_{\xi+1}$ and that $t_{\zeta+1} = t'_{\xi+1} + t''_{\eta+1}$, thus completing the proof of the theorem.

An immediate consequence of the "Theorem on the Product" is the

"THEOREM ON COMMON DIVISORS." *If the parentheses* $A(\alpha_0 \neq -\infty)$ *and* $B(\beta_0 \neq -\infty)$ *are such that no* $m'_\xi$ $\{\xi = 1, 2, \cdots, f\}$ *is equal to an* $m''_\eta$ $\{\eta = 1, 2, \cdots, g\}$, *they have no common factors of order* $\geqq 1$.

For every $m$ of the characteristic set of a common factor is an $m$ of both $S_A$ and $S_B$.

## 7. THE "FACTORIZATION THEOREM" AND THE "CRITERION FOR IRREDUCIBILITY"

If the parenthesis $C$ is such that $\gamma_\nu$ $\{\nu = 0, 1, \cdots, n\}$ is an integer, which is the case in nearly all of the applications, $m_\zeta$ $\{\zeta = 1, 2, \cdots, h\}$ is a rational number, being equal to $(\gamma_{s_\zeta} - \gamma_{s_{\zeta-1}})/t_\zeta$ or $-1$ by 1*d*, § 4. We then introduce the following

*Definition of* $\tau_\zeta$ *and of* $d_\zeta$. The number $\tau_\zeta$ is the (positive) denominator of $m_\zeta$ when $m_\zeta$ is written in its lowest terms, and $t_\zeta = d_\zeta \tau_\zeta$.

Suppose now that $C$ is decomposed into the product $A \cdot B$ of two parentheses, $A$ being subject to the further restriction that $\alpha_\lambda$ $\{\lambda = 0, 1, \cdots, r\}$ is an integer. ($B$ may be treated as we are going to deal with $A$.) Let $\tau'_\xi$ and $d'_\xi$ $\{\xi = 1, 2, \cdots, f\}$ be defined for $A$ as $\tau_\zeta$ and $d_\zeta$ were for $C$. From the "Theorem on the Product," it follows that every $m'_\xi$ is an $m_\zeta$; hence every $\tau'_\xi$ is a $\tau_\zeta$. The "Theorem on the Product" tells us furthermore, that if $m'_\xi = m_\zeta$, then $t'_\xi \leqq t_\zeta$; hence $d'_\xi \leqq d_\zeta$. Therefore,

$$r = \sum_{\xi=1}^{\xi=f} t'_\xi = \sum_{\xi=1}^{\xi=f} d'_\xi \tau'_\xi = \sum_{\zeta=1}^{\zeta=h} \nu_\zeta \tau_\zeta,$$

where $\nu_\zeta$ is an integer such that $0 \leqq \nu_\zeta \leqq d_\zeta$. We have thus proved the following

"FACTORIZATION THEOREM." *If* $C$ *is such that* $\gamma_\nu$ $\{\nu = 0, 1, \cdots, n\}$ *is an integer (the value* $-\infty$ *being allowed except, of course, for* $\gamma_0$), *then the order of a factor* $A$ *of* $C$, *where* $\alpha_\lambda$ $\{\lambda = 0, 1, \cdots, r\}$ *is likewise an integer (or* $-\infty$), *is of the form*

$$r = \sum_{\zeta=1}^{h} \nu_\zeta \tau_\zeta,$$

*where* $\nu_\zeta$ *is an integer such that* $0 \leqq \nu_\zeta \leqq d_\zeta$ (*and* $\tau_\zeta$, $d_\zeta$ *are defined as above*).

In particular, the "Factorization Theorem" leads to the following

"CRITERION FOR IRREDUCIBILITY." *If the ranks of the elements of* $\mathfrak{S}$ *are all integers (or* $-\infty$), *and* $(\gamma_n - \gamma_0)/n$ *is a fraction in its lowest terms* $> -1$ *and also* $> (\gamma_\nu - \gamma_0)/\nu$ $\{\nu = 1, 2, \cdots, n-1\}$, $C$ *is irreducible.*

For in this case, the characteristic set of $C$ consists of just one pair

$$(m_1, t_1) = (\overline{\gamma_n - \gamma_0/n}, n)$$

and moreover, since $(\gamma_n - \gamma_0)/n$ is in its lowest terms, $\tau_1 = t_1 = n$ and $d_1 = 1$. Hence, the order of a factor $A$ of $C$ is, according to the "Factorization Theorem," of the form

$$r = \nu_1 t_1 = \nu_1 n = 0 \quad \text{or} \quad n.$$

The difficulty of obtaining for $\mathfrak{T}_{-1}$ a more exhaustive criterion than the "Criterion for Irreducibility" may be judged from the following

*Remarks.* There cannot be a theorem for $\mathfrak{T}_{-1}$ that asserts the irreducibility of even a single parenthesis having more than one element in its characteristic set, unless the theorem asserts nothing for at least one of the situations of § 3. For according to a result of Dumas,* based on more general considerations of Hensel,† *every parenthesis of Situation 2 having two or more elements in its characteristic set is reducible;* that is, in Situation 2, $C$ is reducible unless either

$$-1 < (\gamma_n - \gamma_0)/n \geqq (\gamma_\nu - \gamma_0)/\nu \qquad \{\nu = 1, 2, \cdots, n-1\}$$

or

$$S_C = \{(m_1, t_1)\} = \{(-1, n)\}.$$

Furthermore, the "Criterion for Irreducibility" becomes false, if the condition that $(\gamma_n - \gamma_0)/n$ be a fraction in its lowest terms is not demanded. In fact, for every $n$ and $\gamma_n - \gamma_0$ that are not relatively prime, there is a $C$ belonging to Situation 1, for instance, that satisfies every other condition of our criterion and is nevertheless reducible, as the following example will show: Let $n = kd$ and $\gamma_n - \gamma_0 = ld$, $k$, $l$, $d$ being integers such that $d > 1$ (and $l \gtrless 0$). The parenthesis $C = y^{kd} - p^{-ld}$ of Situation 1 has the factor $y^k - p^{-l}$, although $\gamma_0 = 0$, $\gamma_1 = -\infty$, $\gamma_2 = -\infty$, $\cdots$, $\gamma_{n-1} = -\infty$, $\gamma_n = ld$, $(\gamma_n - \gamma_0)/n = (ld/kd) > (\gamma_\nu - \gamma_0)/\nu$ $\{\nu = 1, 2, \cdots, n-1\}$. Finally, our criterion becomes false if the condition $(\gamma_n - \gamma_0)/n > -1$ is made milder to the extent of replacing 1 by a number $\epsilon > 1$, however small $\epsilon - 1$ may be. (If $(\gamma_n - \gamma_0)/n = -1$, $n$ must be 1 on account of the relative primeness of $n$ and $\gamma_n - \gamma_0$, and our criterion is trivial.) This is shown by the following example in Situation 6:‡ Let

$$C = \frac{d^n y}{dx^n} + x^{-n}\frac{dy}{dx} - nx^{-n-1}y.$$

Here $\gamma_0 = 0$, $\gamma_1 = -\infty$, $\gamma_2 = -\infty$, $\cdots$, $\gamma_{n-2} = -\infty$, $\gamma_{n-1} = -n$, $\gamma_n = -n-1$; $(\gamma_n - \gamma_0)/n = -(n+1)/n > (\gamma_\nu - \gamma_0)/\nu$ $\{\nu = 1, 2,$

* L. c., p. 223.

† *Neue Grundlagen der Arithmetik*, Journal für Mathematik, vol. 127 (1904), pp. 51–84, Section 4.

‡ It will appear in the next section that no such example is possible for Situations 1, 2, 3, 4, 5, 8, or 9 of § 3, since these situations (as will appear) are instances of the Theory $\mathfrak{T}_{-\infty}$. As a consequence, the "Criterion for Irreducibility" holds for these situations if $(\gamma_n - \gamma_0)/n > -1$ is replaced by $(\gamma_n - \gamma_0)/n > -\infty$.

$\cdots, n-1\}$. However small $\epsilon - 1$ may be, $n$ may be so chosen that

$$\frac{\gamma_n - \gamma_0}{n} > -\epsilon.$$

Still, $C = A \cdot B$, where

$$A = dy/dx \qquad \text{and} \qquad B = d^{n-1}y/dx^{n-1} + x^{-n}y.$$

Clearly the "Criterion for Irreducibility" is not improved by replacing $(\gamma_n - \gamma_0)/n > (\gamma_\nu - \gamma_0)/\nu$ by $(\gamma_n - \gamma_0)/n \geqq (\gamma_\nu - \gamma_0)/\nu$; for the former condition is a consequence of the latter and the condition that $(\gamma_n - \gamma_0)/n$ is a fraction in its lowest terms.

## 8. The abstract theory $\mathfrak{T}_\epsilon$

In his 1900 Mathematische Annalen paper, *Ueber die Irreductibilität algebraischer Functionalgleichungen*, Koenigsberger deals with a situation, which, if slightly modified to conform to our point of view, does not fall under the theory $\mathfrak{T}_{-1}$. This situation, which we designate in accordance with our previous numbering as Situation 10, is essentially* as follows:

*Situation* 10. $\mathfrak{S}$ consists of the collectivity of fractions $e = e'/e''$, where $e'$ and $e''$ are power series in $x - \theta$ $(= x_1$, cf. Situation 5) with arbitrary complex coefficients, $\theta$ being constant with reference to $x$. As in Situation 5, the rank of $e\,(\neq 0)$ is defined as $d'' - d'$, where $d'$ and $d''$ represent the lowest degrees (in the usual sense) in $x - \theta = x_1$ of a non-zero term of $e'$ and $e''$ respectively; the rank of 0 is $-\infty$. The parenthesis $(e_0, e_1, \cdots, e_m)$ is the linear homogeneous differential expression

$$e_0(x)\frac{d^m y}{dx^m} + e_1(x)\frac{d^{m-1}y}{dx^{m-1}} + \cdots + e_m(x)y.$$

The product of two parentheses is defined as in Situation 6.

If we compare Situation 10 with the illustrative example of § 1 (that is, with Situation 6), we see it to be no longer true, as it was there, that the rank of $de(x)/dx$ is always less than or equal to the rank of $e(x)$; the rank of $x - \theta$, for example, is $-1$, while the rank of $d(x-\theta)/dx$ is 0. It is true, however, as may be verified in elementary fashion, that the rank of $de(x)/dx$ is at most 1 more than the rank of $e(x)$. This new state of affairs requires a corresponding modification of the abstract theory $\mathfrak{T}_{-1}$. A natural modification leads to a theory, called—with propriety in the light of the definition of $\mathfrak{T}_\epsilon$ below—

The Abstract Theory $\mathfrak{T}_{+1}$. *Modify the Theory* $\mathfrak{T}_{-1}$ *by replacing in* (*a*) *and* (*b*), § 2, *the expression* $\alpha_\lambda + \beta_\mu - \sigma$ *by* $\alpha_\lambda + \beta_\mu + \sigma$, *leaving everything else unaltered; the resulting theory is* $\mathfrak{T}_{+1}$.

* Situation 10 is slightly more general than Koenigsberger's situation.

The reader may easily verify that *Situation* 10 *is an instance of the Theory* $\mathfrak{T}_{+1}$.

Again, we proceed to show how another theory, properly called $\mathfrak{T}_{-\infty}$, naturally arises. It is easy to see that most of the situations of § 3 satisfy more restrictive conditions than (*a*) and (*b*) of § 2. In fact, the reader may verify that Situations 1, 2, 3, 4, 5, 8, and 9 actually satisfy the following conditions:

(*a*) $\gamma_\nu \leqq \max_{\lambda+\mu=\nu} (\alpha_\lambda + \beta_\mu)$, $\{0 \leqq \lambda \leqq r, 0 \leqq \mu \leqq s\}$.

(*b*) $\gamma_\nu = \alpha_{\bar\lambda} + \beta_{\bar\mu}$ $\{\bar\lambda + \bar\mu = \nu\}$ if $\alpha_{\bar\lambda} + \beta_{\bar\mu} > \alpha_\lambda + \beta_\mu$ $\{\lambda + \mu = \nu\}$ for every $[\lambda, \mu] \neq [\bar\lambda, \bar\mu]$, $\{0 \leqq \lambda \leqq r, 0 \leqq \mu \leqq s\}$.

It is evident, if we agree that $-\infty \cdot 0 = 0$, that these new conditions (*a*) and (*b*) are equivalent to the following:

(*a*) $\gamma_\nu \leqq \max_{\lambda+\mu+\sigma=\nu} (\alpha_\lambda + \beta_\mu - \infty \cdot \sigma)$, $\{0 \leqq \lambda \leqq r, 0 \leqq \mu \leqq s, 0 \leqq \sigma\}$.

(*b*) $\gamma_\nu = \alpha_{\bar\lambda} + \beta_{\bar\mu}$ $\{\bar\lambda + \bar\mu = \nu\}$ if $\alpha_{\bar\lambda} + \beta_{\bar\mu} > \alpha_\lambda + \beta_\mu - \infty \cdot \sigma$ $\{\lambda + \mu + \sigma = \nu\}$ for every $[\lambda, \mu, \sigma] \neq [\bar\lambda, \bar\mu, 0]$, $\{0 \leqq \lambda \leqq r, 0 \leqq \mu \leqq s, 0 \leqq \sigma\}$.

We thus arrive at

THE ABSTRACT THEORY $\mathfrak{T}_{-\infty}$. *Modify the Theory* $\mathfrak{T}_{-1}$ *by replacing in* (*a*) *and* (*b*), § 2, *the expression* $\alpha_\lambda + \beta_\mu - \sigma$ *by* $\alpha_\lambda + \beta_\mu - \infty \cdot \sigma$ *and agree that* $-\infty \cdot 0 = 0$; *the resulting theory is* $\mathfrak{T}_{-\infty}$.

*Situations* 1–5 *incl.*, 8 *and* 9 *are instances of the Theory* $\mathfrak{T}_{-\infty}$.

The considerations above lead to the idea of

THE ABSTRACT THEORY $\mathfrak{T}_\epsilon$ FOR EVERY REAL NUMBER $\epsilon$ (AND FOR $\epsilon = -\infty$). *This is obtained from* $\mathfrak{T}_{-1}$ *by replacing in* (*a*) *and* (*b*), § 2, *the expression*

$$\alpha_\lambda + \beta_\mu - \sigma \qquad \text{by} \qquad \alpha_\lambda + \beta_\mu + \epsilon\sigma.$$

It is evident that *the Theory* $\mathfrak{T}_\epsilon$ *is a* $\mathfrak{T}_{\epsilon'}$ *for* $\epsilon' > \epsilon$.

In the Theory $\mathfrak{T}_\epsilon$, "characteristic set," the "sum of two characteristic sets," "conjugate segments" and the numbers $\tau_\zeta$ and $d_\zeta$ of § 7 may be defined similarly as in the Theory $\mathfrak{T}_{-1}$.

Thus, *the "characteristic set" of C in the Theory* $\mathfrak{T}_\epsilon$ *is defined as follows*:

$$m_{\zeta+1} = \max_{\nu > s_\zeta} \left( \frac{\gamma_\nu - \gamma_{s_\zeta}}{\nu - s_\zeta}, \epsilon \right).$$

$s_{\zeta+1}$ = *last* $\nu > s_\zeta$ *such that* $\dfrac{\gamma_\nu - \gamma_{s_\zeta}}{\nu - s_\zeta} = m_{\zeta+1}$, *if* $m_{\zeta+1} > \epsilon$; *and*

$s_{\zeta+1} = n$, *if* $m_{\zeta+1} = \epsilon$.

It is readily seen that *the relations* 1, § 4, *hold in* $\mathfrak{T}_\epsilon$ *if the last relation of* (*b*) *is changed to* $m_\lambda \geqq \epsilon$ *and in* (*d*) $m_\zeta > -1$ *is changed to* $m_\zeta > \epsilon$. *Likewise, Lemma I holds if* $m_\zeta > -1$ *is replaced by* $m_\zeta > \epsilon$, and the proof is precisely analogous.

*The "sum of two characteristic sets" and "conjugate segments" are defined in $\mathfrak{T}_\epsilon$ precisely as in § 5. Lemma II holds in $\mathfrak{T}_\epsilon$ if $\Delta$ is set equal to*

$$\alpha_\lambda + \beta_\mu + \epsilon\sigma - (\alpha'_{s_\xi} + \beta''_{s_\eta}),$$

*and in (b) $m'_\xi > -1$, $m''_\eta > -1$ are replaced by $m'_\xi > \epsilon$, $m''_\eta > \epsilon$, and the proof is precisely analogous.*

From these considerations, it is evident that

*The "Theorem on the Product" and the "Theorem on Common Divisors" hold for $\mathfrak{T}_\epsilon$,* and the proof is entirely analogous to that of § 6.

In trying to carry over the results of § 7 to the Theory $\mathfrak{T}_\epsilon$, we at first note a slight divergence; for even if the parenthesis $C$ is such that $\gamma_\nu$ $\{\nu = 0, 1, \cdots, n\}$ is an integer, $m_\zeta$ need not be rational if equal to $\epsilon$. The divergence disappears, however, and in *natural fashion, if we define $\tau_\zeta$ to be equal to $+1$ in case $m_\zeta = \epsilon$, whether $\epsilon$ is rational or not.* In other cases $(m_\zeta > \epsilon)$, $\tau_\zeta$ is defined as in § 7 and $d_\zeta$ (as there) always by the equation $t_\zeta = d_\zeta \tau_\zeta$.* It then appears that

*The "Factorization Theorem" holds in $\mathfrak{T}_\epsilon$* and

*The "Criterion for Irreducibility" holds in $\mathfrak{T}_\epsilon$ if the relation*

$$(\gamma_n - \gamma_0)/n > -1$$

*is replaced by the relation*

$$(\gamma_n - \gamma_0)/n > \epsilon.$$

The remarks at the close of § 7 show the difficulty of obtaining a more exhaustive criterion for irreducibility in $\mathfrak{T}_\epsilon$.

## Appendix. Connections with the literature; concrete examples

(*A*) The Schoenemann-Eisenstein theorem says that

$$C = y^n + i_1 py^{n-1} + i_2 py^{n-2} + \cdots + i_{n-1} py + i_n p,$$

where the $i$'s are integers and $i_n$ not divisible by the prime integer $p$, is irreducible in the domain of rational numbers. This theorem is a special case of the "Criterion for Irreducibility" in the Theory $\mathfrak{T}_{-\infty}$ (or in $\mathfrak{T}_\epsilon$ for $\epsilon < -\frac{1}{2}$), when interpreted for Situation 1, § 3. For, $C = (c_0, c_1, \cdots, c_n)$, where $c_0 = 1$, $c_1 = i_1 p$, $\cdots$, $c_n = i_n p$; hence $\gamma_0 = 0$, $\gamma_1 \leqq -1$, $\gamma_2 \leqq -1$, $\cdots$, $\gamma_{n-1} \leqq -1$, $\gamma_n = -1$. Therefore,

* The reader will remember that in case $m_\zeta = \epsilon$, we have only an inequality

$$(\gamma_{s_\zeta} - \gamma_{s_{\zeta-1}} \leqq m_\zeta t_\zeta)$$

for $\gamma_{s_\zeta}$ instead of the usual equality.

$$(\gamma_n - \gamma_0)/n = -1/n > -\infty, \quad > (\gamma_\nu - \gamma_0)/\nu \quad \{\nu = 1, 2, \cdots, n-1\}$$

and is in its lowest terms.*

The irreducibility of $y^{p-1} + y^{p-2} + \cdots + y + 1$ ($p$ = a prime integer) in the domain of rational numbers is a direct consequence of the Schoenemann-Eisenstein theorem.†

(*B*) The Floquet Theorem‡ says that the linear homogeneous differential expression

$$C = (x-\theta)^{n+1} p_0(x) \frac{d^n y}{dx^n} + (x-\theta)^{n-1+\mu_1} p_1(x) \frac{d^{n-1} y}{dx^{n-1}} + \cdots + (x-\theta)^{1+\mu_{n-1}} p_{n-1}(x) \frac{dy}{dx} + p_n(x) y$$

is irreducible under the following conditions:

1. $C$ is said to be irreducible, if it is not expressible as a product of two linear homogeneous differential expressions of order $\geqq 1$, whose coefficients are power series in $x - \theta$ with arbitrary complex numerical coefficients.

2. $\mu_1, \mu_2, \cdots, \mu_{n-1}$ are all $\geqq 1$.

3. The $p$'s are power series in $x - \theta$ such that $p_0(\theta), p_1(\theta), \cdots, p_n(\theta)$ are all $\neq 0$.

The theorem is a special case of the "Criterion for Irreducibility" in the Theory $\mathfrak{T}_{+1}$, when interpreted for Situation 10, § 8. For, we have

$$\gamma_0 = -(n+1), \quad \gamma_1 = -(n-1+\mu_1), \quad \cdots, \quad \gamma_{n-1} = -(1+\mu_{n-1}), \quad \gamma_n = 0.$$

Hence

$$(\gamma_\nu - \gamma_0)/\nu \ \{\nu = 1, 2, \cdots, n-1\} = (\nu + 1 - \mu_\nu)/\nu \leqq 1 < (n+1)/n = (\gamma_n - \gamma_0)/n,$$

a fraction in its lowest terms.

(*C*) Koenigsberger, Journal für Mathematik, vol. 115, pp. 53–78. Every result of this paper (pertaining exclusively to the reducibility of expressions) falls under our theory. We select several examples.§

(*a*) P. 55. "Wenn in einer algebraischen Gleichung

$$C = c_0(x) y^n + c_1(x) y^{n-1} + \cdots + c_{n-1}(x) y + c_n(x) = 0,$$

* By noting that $C$ is irreducible when and only when $(c_n, c_{n-1}, \cdots, c_0)$ is irreducible—this is, of course, not always so in the abstract theory—we may show that the Schoenemann-Eisenstein theorem is a consequence of the "Criterion for Irreducibility" in a $\mathfrak{T}_\epsilon$, $\epsilon \leqq 0$, interpreted for Situation 1, Section 3. A similar remark applies to other examples of the appendix.

† Eisenstein, l. c.

‡ Contained in Koenigsberger, Mathematische Annalen, vol. 53 (1900), p. 67

§ In the quoted theorems of the appendix, slight deviations from the original are made when greater harmony in statement with the rest of this paper may be thereby secured.

in welcher $c_0(x)$, $c_1(x)$, $\cdots$, $c_n(x)$ ganze Functionen* von $x$ bedeuten, alle Coefficienten mit Ausnahme des ersten durch einen Linearfactor $x - \theta$, der letzte aber nicht durch $(x-\theta)^2$ theilbar ist, so ist die Gleichung irreductibel."

This theorem† is a special case of the "Criterion for Irreducibility," $\mathfrak{T}_{-\infty}$, Situation 5, § 3, $k$ being taken equal to 1, and $x_1$ set equal to $x - \theta$. It is seen that $\gamma_0 = 0, \gamma_1 \leqq -1, \gamma_2 \leqq -1, \cdots, \gamma_{n-1} \leqq -1, \gamma_n = 1$, as in $(A)$.

(*b*) P. 63. THEOREM. Every polynomial of the form

$$C = p_0(x)y^n + (x-\theta)^{E(i/n)+1} p_1(x) y^{n-1} + (x-\theta)^{E(2i/n)+1} p_2(x) y^{n-2} + \cdots + (x-\theta)^{E[(n-1)i/n]+1} p_{n-1}(x) + (x-\theta)^i p_n(x)$$

is irreducible under the following conditions:

1. The $p$'s are polynomials in $x$, and the domain of rationality is the set of rational functions of $x$ with arbitrary complex coefficients; $p_0(\theta) \neq 0$ and $p_n(\theta) \neq 0$.

2. $i$ is a positive integer relatively prime to $n$; $E(t)$ signifies, as usual, the largest integer $\leqq t$.

The theorem is a special case of the "Criterion for Irreducibility" in $\mathfrak{T}_{-\infty}$, Situation 5, $k = 1$, $x_1 = x - \theta$. For, if $\gamma_\nu$ represents the rank of the coefficient of $y^{n-\nu}$, we have $\gamma_0 = 0, \gamma_1 \leqq -E(i/n) - 1, \gamma_2 \leqq -E(2i/n) - 1, \cdots, \gamma_{n-1} \leqq -E(\overline{n-1}i/n) - 1, \gamma_n = -i$; $(\gamma_n - \gamma_0)/n = -i/n$, a fraction in its lowest terms and furthermore $(\gamma_n - \gamma_0)/n > (\gamma_\nu - \gamma_0)/\nu$ $\{\nu = 1, 2, \cdots, n-1\}$.

(*c*) P. 77. "Jede algebraische Gleichung $n$-ten Grades von der Form

$$C = i_0 y^n + p^{\kappa_\mu^{(1)}+(1/\mu)} q^{\kappa_\nu^{(1)}+(1/\nu)} i_1 y^{n-1} + p^{\kappa_\mu^{(2)}+(2/\mu)} q^{\kappa_\nu^{(2)}+(2/\nu)} i_2 y^{n-2} + \cdots + p^{\kappa_\mu^{(n-1)}+(\overline{n-1}/\mu)} q^{\kappa_\nu^{(n-1)}+(\overline{n-1}/\nu)} i_{n-1} y + p^{\kappa_\mu^{(n)}} q^{\kappa_\nu^{(n)}} i_n = 0,$$

in welcher $n = \mu \cdot \nu$, $\mu$, und $\nu$ relative Primzahlen, $p$ und $q$ zwei verschiedene Primzahlen, $i_0, i_1, \cdots, i_n$ beliebige ganze Zahlen, von denen nur $i_0$ und $i_n$ weder durch $p$ noch durch $q$ theilbar sind,

$$\kappa_\mu^{(\rho)} = E(\overline{\rho - 1}/\mu) + 1, \qquad \kappa_\nu^{(\rho)} = E(\overline{\rho - 1}/\nu) + 1,$$

also $\kappa_\mu^{(n)} = \nu$, $\kappa_\nu^{(n)} = \mu$ ist, endlich das Symbol $(\rho/\sigma)$ die Einheit oder Null bedeutet, je nachdem $\rho$ durch $\sigma$ theilbar oder nicht theilbar, ist irreductibel."

The theorem follows from the "Factorization Theorem," $\mathfrak{T}_{-\infty}$, Situation 1. For, let $\gamma_\nu$ be the rank of the coefficient of $y^{n-\nu}$ with respect to the prime $p$;

* By "ganze Functionen," Koenigsberger means polynomials; the theorem is also true, however, when the $c$'s are power series in $x - \theta$.

† The domain of rationality is not given explicitly in the statement of the theorem; however, the subsequent statement that the theorem comes under Situation 5 supplies the deficiency. A similar remark applies to other examples of the appendix.

$\gamma'_\nu$, the rank of the same coefficient with respect to the prime $q$. We then have:

$$\gamma_0 = 0, \quad \gamma_1 \leqq -\kappa_\mu^{(1)} - (1/\mu), \quad \gamma_2 \leqq -\kappa_\mu^{(2)} - (2/\mu), \cdots,$$

$$\gamma_{n-1} \leqq -\kappa_\mu^{(n-1)} - (\overline{n-1}/\mu), \quad \gamma_n = -\kappa_\mu^{(n)} = -\nu;$$

$$(\gamma_n - \gamma_0)/n = -\nu/n \geqq (\gamma_\nu - \gamma_0)/\nu \quad \{\nu = 1, 2, \cdots, n-1\}.$$

The characteristic set (with respect to $p$) consists of the single element $(m_1, t_1) = (-\nu/n, n)$; $\tau_1 = \mu$ and $d_1 = \nu$. Hence, according to the "Factorization Theorem," the order $r$ of a factor of $C$ is of the form $r = i\mu$, where $i$ is an integer. Similarly, by reasoning with the ranks $\gamma'_\nu$ of the coefficients with respect to $q$, we conclude that $r = j\nu$, where $j$ is an integer. Since $\mu$ and $\nu$ are relatively prime, $r$ must be of the form $\kappa\mu\nu = \kappa n$. Hence, $r = 0$ or $n$; that is, $C$ is irreducible.

*Example.*

$$C = i_0 y^6 + pqi_1 y^5 + p^2 qi_2 y^4 + p^2 q^2 i_3 y^3 + p^3 q^2 i_4 y^2 + p^3 q^2 i_5 y + p^3 q^2 i_6 = 0,$$

$i_0$ and $i_6$ being divisible by neither $p$ nor $q$.

It is evident that a more general theorem may be at once obtained by taking $k$ primes ($k$ arbitrary) instead of the two primes $p$ and $q$.

(*D*) Netto, Mathematische Annalen, vol. 48, pp. 81–88. All the results may be treated from our point of view. We select several examples.

(*a*) Pp. 82–83. THEOREM. *The polynomial*

$$y^{2n+1} + i_1 py^{2n} + \cdots + i_n py^{n+1} + i_{n+1} p^2 y^n + \cdots + i_{2n+1} p^2,$$

*where $p$ is a prime, $i_\nu$ $\{\nu = 1, 2, \cdots, 2n+1\}$ an integer and $i_{2n+1}$ not divisible by $p$, is irreducible.*

The theorem is a special case of the "Criterion for Irreducibility," $\mathfrak{T}_{-\infty}$, Situation 1.

(*b*) P. 85. "Wenn im Polynome

$$C = y^n + c_1 y^{n-1} + c_2 y^{n-2} + \cdots + c_{n-1} y + c_n$$

alle Coefficienten $c_\nu$ durch $p$ theilbar sind, $c_{n-1}$ aber durch keine höhere Potenz von $p$ als die erste, dann ist es entweder irreductibel oder es zerfällt in der Weise, dass es einen Factor des Grades 1 und einen anderen irreductiblen Factor des Grades $n-1$ besitzt."

The theorem is a special case of the "Factorization Theorem," $\mathfrak{T}_{-\infty}$, Situation 1. For

$$\gamma_0 = 0, \quad \gamma_1 \leqq -1, \quad \cdots, \quad \gamma_{n-1} = -1, \quad \gamma_n \leqq -1;$$

$$(\gamma_n - \gamma_0)/n \geqq (\gamma_{n-1} - \gamma_0)/(n-1) > (\gamma_\nu - \gamma_0)/\nu$$

$$\{\nu = n-2, n-3, \cdots, 1\}.$$

It is hence evident that the first pair $(m_1, t_1)$ of the characteristic set of $C$ is either $(-1/(n-1), n-1)$—and then $t_2$ of the second pair equals 1—or $(-1/n, n)$. Therefore, according to the "Factorization Theorem," the order of a factor of $C$ is of the form $i(n-1)+j$ or $in$, where $i$ and $j$ have the values 0 or 1. The theorem is thus proved.

(*E*) Perron, Mathematische Annalen, vol. 60, pp. 448–458. Every result falls under our theory. We select several examples.

(*a*) P. 452. "Die algebraische Gleichung

$$C = y^n + p_1^{E(e_1/n)+1} p_2^{E(e_2/n)+1} \cdots p_k^{E(e_k/n)+1} i_1 y^{n-1} + p_1^{E(2e_1/n)+1} \cdots p_k^{E(2e_k/n)-1} i_2 y^{n-2}$$

$$+ \cdots + p_1^{E((n-1)e_1/n)+1} \cdots p_k^{E[(n-1)e_k/n]+1} i_{n-1} y + p_1^{e_1} \cdots p_k^{e_k} i_n = 0,$$

wo $i_1, \cdots, i_n$ beliebige ganze rationale Zahlen bedeuten, deren letzte zu den Primzahlen $p_1, p_2, \cdots, p_k$ prim ist, wo ferner die $k+1$ Zahlen $n, e_1, e_2, \cdots, e_k$ keinen gemeinsamen Teiler haben, ist irreduzibel." (The $e$'s are $>0$.)

The theorem is a special case of the "Factorization Theorem," $\mathfrak{T}_{-\infty}$, Situation 1. For, let $\gamma_{\nu,\sigma}$ $\{\nu = 0, 1, \cdots, n;\ \sigma = 1, 2, \cdots, k\}$ be the rank of the coefficient of $y^{n-\nu}$ with respect to the prime $p_\sigma$. Then

$$\gamma_{0,\sigma} = 0, \quad \gamma_{1,\sigma} \leqq -(e_\sigma/n) - \theta_{1,\sigma}, \quad \gamma_{2,\sigma} \leqq -(2e_\sigma/n) - \theta_{2,\sigma}, \quad \cdots,$$

$$\gamma_{n-1,\sigma} \leqq -((n-1)e_\sigma/n) - \theta_{n-1,\sigma}, \qquad \gamma_{n,\sigma} = -e_\sigma,$$

where the $\theta$'s lie between 0 (exclusive) and 1 (inclusive);

$$(\gamma_{n,\sigma} - \gamma_{0,\sigma})/n = -(e_\sigma/n) > (\gamma_{\nu,\sigma} - \gamma_{0,\sigma})/\nu \quad \{\nu = 1, 2, \cdots, n-1\}.$$

It then follows from the "Factorization Theorem," the ranks of the coefficients being taken with respect to $p_\sigma$, that the characteristic set consists of the single element $(-e_\sigma/n, n)$. Hence, the order $r$ of a factor of $C$ has the form $(in/[n, e_\sigma])$, where $i$ is an integer and $[n, e_\sigma] = G.C.D.$ of $n$ and $e_\sigma$. Therefore $r[n, e_\sigma]$ is a multiple of $n$ for $\sigma = 1, 2, \cdots, k$. Since the *G.C.D.* of the $k$ numbers $[n, e_\sigma]$ is 1 by hypothesis, there exist $k$ integers $j_\sigma$ such that $\sum_{\sigma=1}^{\sigma=k} j_\sigma [n, e_\sigma] = 1$. Hence $r = \sum_{\sigma=1}^{\sigma=k} j_\sigma r[n, e_\sigma]$ is a multiple of $n$; that is, $r = 0$ or $n$.

(*b*) Pp. 455–456. "Das Polynom

$$y^n + i_1 y^{n-1} + \cdots + i_k y^{n-k} + p i_{k+1} y^{n-k-1} + \cdots + p i_n,$$

oder allgemeiner

$$C = y^n + i_1 y^{n-1} + \cdots + i_k y^{n-k} + p^{E(e/\overline{n-k})+1} i_{k+1} y^{n-k-1} + \cdots$$
$$+ p^{E[(n-k-1)e/\overline{n-k}]+1} i_{n-1} y + p^e i_n,$$

wo $i_k$ und $i_n$ durch die Primzahl $p$ nicht teilbar sind, und wo $e$ zu $n-k$ relativ prim ist, enthält mindestens *einen* irreduzibeln Faktor, dessen Grad $\geqq n-k$ ist."

A special case of the "Factorization Theorem," $\mathfrak{T}_{-\infty}$, Situation 1.

(*F*) Koenigsberger, Mathematische Annalen, vol. 53, pp. 49–80. Every result (within our present field of investigation) is a special case of the "Factorization Theorem" either in $\mathfrak{T}_{+1}$ for Situation 10 or in $\mathfrak{T}_{-\infty}$ for Situation 5, $k=1$, $x_1 = x-\theta$. We select several illustrations.

(*a*) P. 69. "Die lineare Differentialgleichung

$$C = (x-\theta)^{n+2} p_0(x) \frac{d^n y}{dx^n} + (x-\theta)^{n-1+\mu_1} p_1(x) \frac{d^{n-1} y}{dx^{n-1}} + \cdots$$
$$+ (x-\theta)^{1+\mu_{n-1}} p_{n-1}(x) \frac{dy}{dx} + p_n(x) y = 0,$$

in welcher $\mu_1, \mu_2, \cdots, \mu_{n-1} \geqq 2$ sind, ist, wenn die Ordnung $n$ derselben ungerade, stets irreductibel; ist die Ordnung jedoch gerade, so kann sie nur mit einer gleichartigen linearen Differentialgleichung $n/2$ ter Ordnung die sämmtlichen Integrale der letzteren gemein haben." (Here the $p$'s are power series in $(x-\theta)$ such that $p_0(\theta), \cdots, p_n(\theta)$ are all $\neq 0$; cf. (*B*), appendix.)

This theorem falls in $\mathfrak{T}_{+1}$, Situation 10. We have

$$\gamma_0 = -(n+2), \quad \cdots, \quad \gamma_\nu = -(n-\nu+\mu_\nu), \quad \cdots, \quad \gamma_n = 0;$$
$$(\gamma_n - \gamma_0)/n = (n+2)/n > (\nu+2-\mu_\nu)/\nu = (\gamma_\nu - \gamma_0)/\nu$$
$$\{\nu = 1, 2, \cdots, n-1\}.$$

Hence, according to the "Criterion for Irreducibility," $C$ is irreducible if $n$ is odd; and according to the "Factorization Theorem," if $n$ is even, the order of a factor of $C$ is of the form $in/2$, where $i$ is an integer. This result is equivalent to the statement of the theorem.

(*b*) P. 73. "Eine Differentialgleichung von der Form

$$(x-\theta)^{n+\mu_0} p_0(x) \frac{d^n y}{dx^n} + (x-\theta)^{n-1+\mu_1} p_1(x) \frac{d^{n-1} y}{dx^{n-1}} + \cdots$$
$$+ (x-\theta)^{1+\mu_{n-1}} p_{n-1}(x) \frac{dy}{dx} + p_n(x) y = 0,$$

in welcher $\mu_1, \mu_2, \cdots, \mu_{n-1} \geqq \mu_0$ sind, kann nur mit einer gleichartigen Differentialgleichung, deren Ordnung $\geqq n/\mu_0$ ist, alle Integrale der letzteren gemein haben." (Here the $p$'s are as in (*a*).)

A special case of the " Factorization Theorem," $\mathfrak{T}_{+1}$, Situation 10.

(*c*) P. 77. " Die lineare Differentialgleichung

$$(x-\theta)^{\rho n+1} p_0(x)\frac{d^n y}{dx^n} + (x-\theta)^{\rho(n-1)+\nu_1} p_1(x)\frac{d^{n-1} y}{dx^{n-1}} + (x-\theta)^{\rho(n-2)+\nu_2} p_2(x)\frac{d^{n-2} y}{dx^{n-2}} + \cdots + (x-\theta)^{\rho+\nu_{n-1}} p_{n-1}(x)\frac{dy}{dx} + p_n(x)y = 0,$$

in welcher $\nu_1, \nu_2, \cdots, \nu_{n-1} \geqq 1$ und $\rho$ eine beliebige positive ganze Zahl bedeutet, ist irreductibel." (Here the $p$'s are as in (*a*).)

A consequence of the " Criterion for Irreducibility," $\mathfrak{T}_{+1}$, Situation 10.

(*G*) Koenigsberger, Journal für Mathematik, vol. 121, pp. 320–359. The leading results on factorization contained in this paper—in particular, the (complicated but general) result p. 342—fall under the " Factorization Theorem," $\mathfrak{T}_{-\infty}$, either for Situation 1 or for Situation 5, $k = 1$.

(*H*) Perron's Criterion, Journal für Mathematik, vol. 132, p. 304.

" Wenn sich unter den respektiven Gradzahlen $\gamma_1, \gamma_2, \cdots, \gamma_n$ der Funktionen $c_1(x), c_2(x), \cdots, c_n(x)$ eine findet $(\gamma_i)$, welche den Ungleichungen

$$(\alpha) \quad \gamma_i > \gamma_\nu$$

$$(\beta) \quad \frac{\gamma_i}{i} > \frac{\gamma_\nu}{\nu}$$

genügt, wo $\nu$ jede von $i$ verschiedene der Zahlen $1, 2, \cdots, n$ bedeutet; und wenn ausserdem $\gamma_i$ relativ prim ist zu $i$, so ist

$$C = f(x, y) = y^n + c_1(x)y^{n-1} + \cdots + c_n(x)$$

irreduzibel. Sind aber die Ungleichungen $(\alpha)$ nur insoweit erfüllt, dass für gewisse Werte von $\nu$ das Gleichheiszeichen an Stelle des Grösserzeichens tritt, so ist für die Irreduzibilität ausserdem noch erforderlich, dass keine Konstante $c$ existiert, für die $f(x, c)$ als Funktion von $x$ identisch verschwindet." (The $c$'s are here polynomials in $x$ with arbitrary complex coefficients.)

This theorem is a consequence of the " Theorem on the Product," $\mathfrak{T}_{-\infty}$, Situation 3, $k = 1$, $x_1 = x$. For, from $(\beta)$ it follows, since $\gamma_0 = 0$, that the first element $(m_1, t_1)$ of the characteristic set of $C$ is $(\gamma_i/i, i)$. If $i = n$, it follows from the " Criterion for Irreducibility," since $\gamma_i$ and $i$ are relatively prime, that $C$ is irreducible. If $i < n$, the characteristic set of $C$ has other elements $(m_\zeta, t_\zeta)$, where $\zeta > 1$, but on account of $(\alpha)$, $m_\zeta$ must be $< 0$. Now, according to the " Theorem on the Product " [cf. proof of the " Factorization Theorem "], it follows in virtue of the relative primeness of $\gamma_i$

and $i$, that only *one* of the factors, say $A$, of $C = A \cdot B$ ($r \geqq 1$ and $s \geqq 1$) can have $m_i = \gamma_i/i$ as an $m$ of its characteristic set; and every $m$ of the characteristic set of $B$ is thus $< 0$. According to a theorem of Gauss, we may assume, without loss of generality, that the coefficients of the various powers of $y$ in $B$ are *polynomials* in $x$ and hence of rank $\geqq 0$. Since $\beta_0$ ( = rank of the coefficient of the highest power of $y$ in $B$) is then obviously 0, at least one $m$ of the characteristic set of $B$ is $\geqq 0$. This contradiction shows that $C$ is irreducible. If, on the other hand, the inequality ($\alpha$) is changed to $\gamma_i \geqq \gamma_\nu$, we have no longer a contradiction, but can merely conclude that every $m$ of the characteristic set of $B$ is 0. Hence the rank of every coefficient of the powers of $y$ in $B$ is 0; that is, these coefficients are constants. Therefore, $B$ depends on $y$ alone, and there exists a value $c$ for $y$, namely a root for $y$ of $B = 0$, such that $C = f(x, c) = A \cdot B \equiv 0$.

(*I*) Dumas, Journal de Mathématiques, ser. 6, vol. 2, pp. 191–258. The theorem p. 217 is equivalent to the "Theorem on the Product"; the theorem p. 237, to the "Factorization Theorem"; and the theorem p. 242, to the "Theorem on Common Divisors," all for $\mathfrak{T}_{-\infty}$, Situation 2.

(*J*) Bauer, Journal für Mathematik (see previous citations). The principal results come under the results of Dumas.

University of Nebraska,
*Dec.* 30, 1915.

# THE FORMAL MODULAR INVARIANT THEORY OF BINARY QUANTICS*

BY

O. E. GLENN

The group, $G$, represented by the general binary linear transformation in which the coefficients are parameters representing residues of the prime number $p$, has the universal covariants†

$$L_t = x_1^{p^t} x_2 - x_1 x_2^{p^t},$$

a fundamental system being given by $L = L_1$, $Q = L_2/L_1$. Assume that

$$f_m = (a_0, a_1, \cdots, a_m)(x_1, x_2)^m$$

is a binary quantic of order $m$ whose coefficients are variables. We study the functions of the $a$'s and $x$'s which are invariantive under the operation of transforming $f_m$ by $G$. Sections 1, 2, 3 deal with methods of constructing such functions, especial attention being given to the case $p = 2$. In § 5 fundamental systems of first degree concomitants of the quartic and quintic are derived. Section 4 is devoted to a proof that if the systems of concomitants modulo 2 of the forms of orders 1, 2, 3 and the simultaneous systems obtained by combining forms of these three orders are all finite, then the system of $f_m$ is likewise finite. The hitherto undemonstrated theorem on the finiteness of the formal concomitants modulo 2 is thus reduced to a comparatively simple problem. In § 6 there is proved a theorem on the reducibility of any covariant, modulo 2, of the binary cubic in terms of a set of fourteen invariants and covariants.

## 1. Invariant operators

In addition to modular polars and transvectants previously discussed‡ by the present writer we cite, in order to augment the number of construction

* Presented to the Society, February 26, 1916.

† Dickson, these Transactions, vol. 12 (1911), p. 75; and *Madison Colloquium Lectures*, 1913, p. 33.

‡ O. E. Glenn, American Journal of Mathematics, vol. 37 (1915), p. 73; and Bulletin of the American Mathematical Society, vol. 21 (1915), p. 167.

methods, the universal operators obtained by replacing, in $L$ and $Q$, the variables $x_1$, $x_2$ by the cogredient elements $\partial/\partial x_2$, $-\partial/\partial x_1$. Call these operators respectively $L_\delta$ and $Q_\delta$;

$$Q_\delta = \frac{\partial^{p(p-1)}}{\partial x_1^{p(p-1)}} + \frac{\partial^{p(p-1)}}{\partial x_1^{(p-1)(p-1)}\,\partial x_2^{p-1}} + \cdots.$$

If $K_M$ is any formal covariant modulo $p$ of degree-order $(i, M)$ and if

(1) $$(p+1)r + p(p-1)s + w = M,$$

then

$$C_{rsw} = L_\delta^r\, Q_\delta^s\, K_M$$

is a formal covariant of degree-order $(i, w)$. The number of concomitants yielded by this formula equals the number of distinct solutions in positive integers $(r, s, w)$ of the linear diophantine equation (1). The lists given below are for $p = 2$, $K_M \equiv f_m$, each giving all forms $C_{rsw}$ for the corresponding $m$. Corresponding to each invariant $C_{rs0}$ there exists an invariant operator $\sum(\partial/\partial a)$ obtained by constructing the Aronhold operator for the two forms $L^r\, Q^s$, $f_m$. These are also given in the particular cases shown.

$$m = 4$$

$$C_{020} = a_1 + a_2 + a_3; \qquad \frac{\partial}{\partial a_0} + \frac{\partial}{\partial a_2} + \frac{\partial}{\partial a_4},$$

$$C_{101} = a_1 x_1 + a_3 x_2, \qquad C_{012} = a_1 x_1^2 + a_3 x_2^2.$$

$$m = 5$$

$$C_{110} = a_1 + a_2 + a_3 + a_4; \qquad \frac{\partial}{\partial a_1} + \frac{\partial}{\partial a_4},$$

$$C_{021} = a_2 x_1 + a_3 x_2, \qquad C_{102} = a_2 x_1^2 + a_3 x_2^2, \qquad C_{013} = QC_{021}.$$

$$m = 6$$

$$C_{200} = C_{030} = a_3; \qquad \frac{\partial}{\partial a_0} + \frac{\partial}{\partial a_1} + \frac{\partial}{\partial a_3} + \frac{\partial}{\partial a_5} + \frac{\partial}{\partial a_6}; \qquad \frac{\partial}{\partial a_2} + \frac{\partial}{\partial a_4},$$

$$C_{111} = (a_1 + a_3)x_1 + (a_3 + a_5)x_2, \qquad C_{022} = a_3 Q, \qquad C_{103} = a_3 L,$$

$$C_{014} = a_1 x_1^4 + a_3 x_1^2 x_2^2 + a_5 x_2^4.$$

$$m = 7$$

$$C_{120} = a_1 + a_2 + a_3 + a_4 + a_5 + a_6; \qquad \sum_{i=1}^{6} \frac{\partial}{\partial a_i},$$

$$C_{201} = (a_2 + a_4)x_1 + (a_3 + a_5)x_2,$$

$$C_{031} = (a_0 + a_1 + a_3 + a_5 + a_6)x_1 + (a_1 + a_2 + a_4 + a_6 + a_7)x_2,$$

$$C_{112} = (a_1 + a_4)x_1^2 + (a_3 + a_6)x_2^2,$$

$$C_{023} = (a_0 + a_2 + a_4)x_1^3 + (a_1 + a_3 + a_5)x_1^2 x_2 + (a_2 + a_4 + a_6)x_1 x_2^2 + (a_3 + a_5 + a_7)x_2^3,$$

$$C_{104} = (a_1 + a_2)x_1^4 + (a_5 + a_6)x_2^4,$$

$$C_{015} = (a_0 + a_1 + a_2)x_1^5 + (a_1 + a_2 + a_3)x_1^4 x_2 + (a_4 + a_5 + a_6)x_1 x_2^4 + (a_5 + a_6 + a_7)x_2^5.$$

## 2. Covariants led by assigned seminvariants

If we add to the variables $a_0, a_1, \cdots; x_1, x_2$ in any formal covariant $\phi(a_0, \cdots; x_1, x_2)$ of $f_m$ the increments of these variables when $f_m$ is transformed by $x_1 = x_1'$, $x_2 = tx_1' + x_2'$, and then expand $\phi(a_0 + \delta a_0, \cdots)$ by Taylor's theorem, we find that any formal covariant modulo $p$ has an annihilator of the type*

$$\Upsilon = O_0 + O_1 x_1 \frac{\partial}{\partial x_2} + O_2 x_1^2 \frac{\partial^2}{\partial x_2^2} + \cdots + O_j x_1^j \frac{\partial^j}{\partial x_2^j} + \cdots.$$

In this theory $t$ is any residue modulo $p$ and the expansion in question contains only the $p$ terms obtained by reducing all powers of $t$ below the $p$th by Fermat's theorem. The operator $O_j$ $(j = 0, 1, \cdots)$ is a partial differential operator in the derivatives with respect to the coefficients of $f_m$, non-homogeneous as to the derivatives the orders of which range from zero to infinity in each $O_j$.

If we apply $\Upsilon$ to a covariant and proceed as in the well-known proof of Roberts' theorem on the unique determination of an algebraical covariant from its seminvariant leader we find that the resulting relations in the covariant's coefficients and the operators $O_j$ $(j = 0, 1, \cdots)$ are not recurrent in the formal theory, whereas they are recurrent in the algebraic theory. Thus a formal covariant is not uniquely determined from its leader. This is also evident from the fact that if $T$ is any such covariant of order $w \not\equiv 0 \pmod{p}$, then

$$\left(x_1^p \frac{\partial}{\partial x_1} + x_2^p \frac{\partial}{\partial x_2}\right) T$$

is also a formal covariant having the same seminvariant leader.

Under definite conditions we can, however, derive a covariant $T$ of order $p - 1$ of $f_m$ having a given seminvariant $S$ as leader. This is done by substituting for $a_0, a_1, a_2, \cdots$ in $S$ the derivatives

* Cf. Dickson, these Transactions, vol. 8 (1907), p. 209. An explicit $\Omega_0(\delta)$, analogous to $O_0$, is given in my paper in American Journal of Mathematics, loc. cit., p. 75.

$$f_m, \quad \frac{1}{m}\frac{\partial f_m}{\partial x_2}, \quad \frac{1}{m(m-1)}\frac{\partial^2 f_m}{\partial x_2^2}, \cdots,$$

regarding $x_1$, $x_2$ as integers modulo $p$ and rearranging the result as a polynomial in $x_1$, $x_2$ of order $p-1$. This is analogous to the process of Faà di Bruno for algebraic covariants. The same result is reached by expanding $S(a_0', a_1', \cdots)$, corresponding to the transformation of $f_m$ by

$$x_1 = x_1', \qquad x_2 = tx_1' + x_2',$$

replacing $t$ by $x_2/x_1$ and multiplying by $x_1^{p-1}$. The result is

$$T = Sx_1^{p-1} + \frac{O_0 S}{\lfloor 1}\, x_1^{p-2} x_2 + \cdots + \frac{O_0^{p-1} S}{\lfloor p-1}\, x_2^{p-1}. \tag{2}$$

This is a formal covariant only provided it satisfies the differential congruence $\Upsilon T \equiv 0 \pmod p$, and is symmetrical under the substitution

$$s = (a_0 a_m)(a_1 a_{m-1}) \cdots (x_1 x_2).$$

Imposing these conditions we deduce the conditions which $S$ must satisfy in order that it may lead the covariant of order $p-1$. These are, ($a$) that $S$ should be transformed into $O_0^{p-1} S/\lfloor p-1$ by $s$, and ($b$) that the following congruences should be satisfied:

$$\begin{aligned}
&O_0^p S \equiv 0,\\
&(O_0^{p-1} + O_1 O_0^{p-1}) S \equiv 0,\\
&(O_0^{p-2} + O_1 O_0^{p-2} + O_2 O_0^{p-1}) S \equiv 0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots \qquad (\mathrm{mod}\ p),\\
&(O_0^2 + O_1 O_0^2 + O_2 O_0^3 + \cdots + O_{p-2} O_0^{p-1}) S \equiv 0,\\
&(O_0 + O_1 O_0 + O_2 O_0^2 + \cdots + O_{p-1} O_0^{p-1}) S \equiv 0.
\end{aligned} \tag{3}$$

The theory of $T$ is thus complicated. It is given here because nothing has been published hitherto on the determination of a covariant with a given leader, and also for the reason that it is of much practical use in constructing covariants with assigned properties in special cases. Let $p = 2$, $m = 3$. A seminvariant of $f_3$ is $S = a_0 a_2 + a_1^2$. Transformation of $f_3$ by $x_1 = x_1'$, $x_2 = tx_1' + x_2'$ gives

$$\begin{aligned}
S' &\equiv [a_0 + t(a_1 + a_2 + a_3)](a_2 + ta_3) + (a_1 + ta_3)^2\\
&\equiv a_0 a_2 + a_1^2 + (a_0 a_3 + a_1 a_2 + a_1 a_3 + a_2^2) t \qquad (\mathrm{mod}\ 2).
\end{aligned}$$

Rendering $S$ symmetrical with the coefficient of $t$ (condition ($a$)) by adding the invariant $a_0 a_3 + a_1 a_2 = \Delta$, (annihilated by $O_0$) we have the formal linear covariant*

*Cf. American Journal of Mathematics, loc. cit., p. 78.

$$C_1^{(2)} = (S + \Delta)x_1 + (\Delta + a_1 a_3 + a_2^2)x_2.$$

For $p = 3$, $m = 2$, $S = a_0 a_1^3 - a_0^3 a_1$, the method leads easily to the quadratic covariant of $f_2$,

$$C_6 = Sx_1^2 + (a_0 a_2^3 - a_0^3 a_2)x_1 x_2 + (a_1 a_2^3 - a_1^3 a_2)x_2^2.$$

### 3. Some covariants of $f_m$ modulo 2

For the modulus 2 the real points $(x_1, x_2)$, other than $(0, 0)$, in the plane are $(1, 1)$, $(0, 1)$, $(1, 0)$. The values of $f_m$ at these points are, respectively,

$$a_0 + a_1 + \cdots + a_m, \qquad a_m, \qquad a_0.$$

By a theorem* of Dickson's, any symmetric function of these three expressions is a formal invariant, modulo 2, all such being, of course, rational in the three elementary symmetric functions

$$K = a_1 + a_2 + \cdots + a_{m-1}, \qquad I = a_0 a_m + (a_0 + a_m)(a_0 + a_m + K),$$
$$k = a_0(a_0 + K + a_m)a_m. \tag{4}$$

In the same way if

$$C_M = A_0 x_1^M + A_1 x_1^{M-1} x_2 + \cdots$$

is any formal covariant of $f_m$, modulo 2, any symmetric function of

$$A_0 + \cdots + A_M, \qquad A_0, \qquad A_M$$

is an invariant of the original form $f_m$. Thus the middle coefficient of any quadratic covariant of $f_m$ is an invariant, the sum of the two middle coefficients of any cubic covariant is an invariant, and so on.

By means of the transformations on the $a$'s induced by $x_1 = x_1' + tx_2'$, $x_2 = x_2'$, viz.

$$a_r' = \left[\binom{m}{r}a_0 + \binom{m-1}{r-1}a_1 + \cdots + \binom{m-r+1}{1}a_{r-1}\right]t + a_r \qquad (r = 0, \cdots, m), \tag{5}$$

it is readily shown that

$$K_2 = a_0 x_1^2 + Kx_1 x_2 + a_m x_2^2, \qquad K_1 = (a_0 + K)x_1 + (K + a_m)x_2, \tag{5$_1$}$$

are formal covariants modulo 2.

Lemma. *For any odd order $m > 1$ $(p = 2)$ the form $f_m$ has a cubic covariant $K_{m3}$ with leading coefficient $a_0$, which is such that*

$$K_{m3} \equiv K_2 \pmod 2$$

*for integral values of $x_1, x_2$.*

* These Transactions, vol. 15 (1914), p. 497.

The covariants $K_{m3}$ for the first twelve odd orders $m$ are

$$K_{33} = a_0 x_1^3 + a_1 x_1^2 x_2 + a_2 x_1 x_2^2 + a_3 x_2^3,$$

$$K_{53} = a_0 x_1^3 + (a_1 + a_2) x_1^2 x_2 + (a_3 + a_4) x_1 x_2^2 + a_5 x_2^3,$$

$$K_{73} = a_0 x_1^3 + (a_1 + a_2 + a_4) x_1^2 x_2 + (a_3 + a_5 + a_6) x_1 x_2^2 + a_7 x_2^3,$$

$$K_{93} = a_0 x_1^3 + (a_1 + a_2 + a_3 + a_4) x_1^2 x_2 + (a_5 + a_6 + a_7 + a_8) x_1 x_2^2 + a_9 x_2^3,$$

$$K_{113} = a_0 x_1^3 + (a_1 + a_2 + a_4 + a_5 + a_8) x_1^2 x_2 + (a_3 + a_6 + a_7 + a_9 + a_{10}) x_1 x_2^2 + a_{11} x_2^3.$$

The general form to be proved is

$$K_{m3} = a_0 x_1^3 + I_1 x_1^2 x_2 + I_2 x_1 x_2^2 + a_m x_2^3,$$

where

$$I_1 = a_{i_1} + \cdots + a_{i_s}, \qquad I_2 = a_{j_1} + \cdots + a_{j_s}, \qquad s = \tfrac{1}{2}(m-1),$$

$$I_1 + I_2 = K.$$

Assuming the increments of $I_1$, $I_2$ under (5) to be $I_1' t$, $I_2' t$ we readily show that

$$I_1' \equiv I_2' \equiv a_0 \pmod 2.$$

Hence, letting $I_1 = \alpha_1 a_1 + \cdots + \alpha_{m-1} a_{m-1}$ ($\alpha_j = 0$ or $1$), we have from $\sum_{r=1}^{r=m-1} \alpha_r a_r'$, with (5), the conclusion that the selection of the set $a_{i_1}, a_{i_2}, \cdots, a_{i_s}$ is accomplished by selecting a set $\sigma_1$ of columns from the following Pascal triangle (6), the set to have the properties ($a$), ($b$) below:

(6)
$$\begin{array}{ccccccccc}
 & & & & & & \binom{2}{1}, & & \\
 & & & & & \binom{3}{1} & \binom{3}{2}, & & \\
 & & & & \cdots & \cdots & \cdots & & \\
 & & & \binom{m-1}{1} & \binom{m-1}{2} & \cdots & \binom{m-1}{m-3} & \binom{m-1}{m-2}, & \\
\binom{m}{1} & \binom{m}{2} & \binom{m}{3} & \cdots & \binom{m}{m-2} & \binom{m}{m-1}, & & & \\
a_1 & a_2 & a_3 & \cdots & a_{m-2} & a_{m-1}. & & &
\end{array}$$

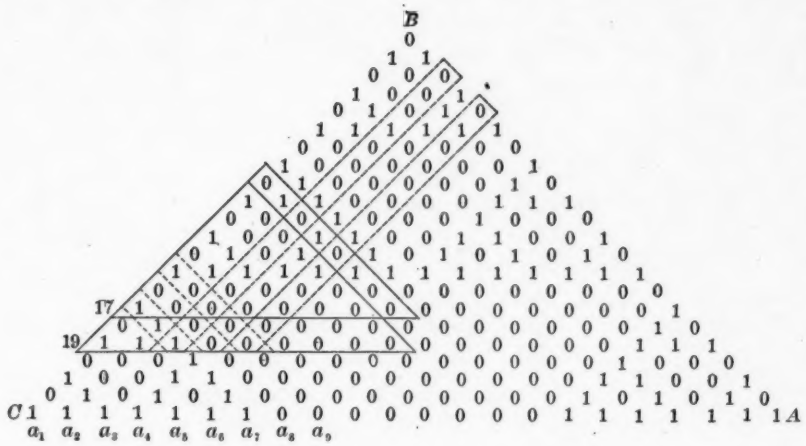
($a$) The sum of the numbers of the first row of the selected set $\sigma_1$ is odd.

($b$) The sum for all of the other rows of $\sigma_1$ is even.

Grant for the moment that such a set can be selected. If its columns be deleted from (6) the set $\sigma_2$ of columns remaining in (6) will also have the properties ($a$), ($b$) since Pascal's triangle is symmetrical with respect to the median drawn from $\binom{2}{1}$. This second set $\sigma_2$ gives $a_{j_1}, a_{j_2}, \cdots, a_{j_s}$, i. e., $I_2$.

We now construct the triangle $ABC$ made up of the residues modulo 2 of

Pascal's triangle, placing $\binom{2}{1}$ at the upper vertex, with the aforesaid median drawn vertically. We note particularly the symmetrical properties, and in particular the regular recurrence of the inverted triangles of zeros, of increasing dimensions, having the median as a line of symmetry. Consider any element $e$ of any horizontal row, to the left of the median. We call the column parallel to $AB$ and containing $e$ the column of $e$. The element $e'$ in the same row as $e$ but occupying the complementary position to the right of the median is called the complementary element of $e$. The column parallel to $AB$ containing $e'$ is the complementary row of $e$, and the column containing $e$ and parallel to $BC$ is the complementary column of $e$. We have $e' = e$, and the complementary row and complementary column of $e$ identical.




Suppose we wish to select $I_1$, i. e., the set $\sigma_1$, for $m = 19$. We take the left-hand half-row number 18 from $B$ as hypotenuse and draw the triangle enclosing the columns of these elements. The elements on the hypotenuse correspond to set

$$T_1 : a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9.$$

If the rows enclosed within this triangle satisfy conditions $(a)$, $(b)$ (as is the case for $m = 17$) then $\sigma_1$ would be the set $T_1$ and $I_1$ equal to the sum of the $\frac{1}{2}(m-1)a$'s in their natural order beginning with $a_1$. The sums of the rows above the first in the triangle are however $1, 1, 0, 1, 0, 1, 0, 1 \pmod 2$ respectively, instead of $0, 0, 0, 0, 0, 0, 0, 0 \pmod 2$ required by condition $(b)$. To secure a triangle (augmented) which does satisfy $(b)$ we replace some of the elements on the hypotenuse by their complementary elements, in this case $e_3, e_5$ (corresponding to $a_3, a_5$) by $e_3', e_5'$ (corresponding to $a_{16}, a_{14}$).

The columns of $e_3$, $e_5$ are thus replaced by their complementary columns by rotating the latter counter-clockwise around $e_3$, $e_5$. The augmented triangle now satisfies conditions $(a)$, $(b)$ and hence

$$I_1 = a_1 + a_2 + a_{16} + a_4 + a_{14} + a_6 + a_7 + a_8 + a_9.$$

Now the orders $m$ for which the triangles do not need to be augmented are those giving the first row below the base of each central triangle of zeros. These values of $m$ are

$$3, 5, 9, 17, 33, \cdots, 2^n + 1, \cdots \tag{7}$$

and $2^n - 1$ is the number of the row, counting downward from $B$, in which the base of the $n$th central triangle of zeros is found. For $m = 2^n + 1$, $I_1$ equals the sum of the first $2^{n-1}$ $a$'s in their natural order, $I_1 = a_1 + a_2 + \cdots$. Between any two consecutive cases of (7) there is a cycle of augmented triangles such that the whole configuration from $B$ downward forms a sequence of recurring cycles of figures proceeding according to the laws of symmetry of the triangle $ABC$. Thus the set $\sigma_1$ can always be selected and the lemma is true.

We now observe that any covariant $C_M$ of $f_m$ of odd order $M$ has corresponding to it a quadratic and a cubic covariant of $f_m$, constructed as covariants of $C_M$, on the models of $K_2$, $K_{m3}$ respectively. (Cf. § 6.)

## 4. The finiteness of the formal concomitants for the order $m$ and modulus 2

Theorem. *The general quantic $f_m$, $m > 3$, is reducible, modulo 2, in terms of $Q$, $L$, and its own covariants of the first degree and orders 1, 2, 3.*

Let $m$ be even, $m = 2k$. Then $f_m - Q^{k-1} K_2$ has the factor $x_2$. Hence it has the factor $L$, since the real points (mod 2) form a conjugate set under the group $G$. Therefore

$$f_m \equiv Q^{k-1} K_2 + LC \pmod 2, \tag{8}$$

where $C$ is a first degree covariant of $f_m$ of odd order $2k - 3$. Next let $m$ be odd, $m = 2h + 3 > 3$. Then

$$f_m \equiv Q^h K_{m3} + LC' \pmod 2, \tag{9}$$

where $C'$ is of even order $2h$.

The forms $C$, $C'$ are, in turn, reducible in terms of their own covariants of orders 2, 3 with $L$ and $Q$, and, as a covariant of $C$ is a covariant of $f_m$, we are led to a recurring process by which $f_m$ is in all cases expressed in terms of its covariants of orders 1, 2, 3, and $L$ and $Q$, which was to be proved.

Any concomitant of a polynomial in a set of concomitants is a function of

concomitants of the forms in the set. Hence if the systems for the orders 1, 2, 3, and the simultaneous systems modulo 2 for these three orders are finite, the system for $f_m$ is likewise finite (see § 6).

## 5. Systems of the first degree

While illustrating formulas (8), (9), we now derive fundamental systems of first degree concomitants for the quartic and the quintic forms.

An independent set of linear seminvariants of $f_4$ is $a_0$, $a_1$, $a_2 + a_3$. The only linear invariant is $C_{020} = a_1 + a_2 + a_3$ ($= K$, $m = 4$ (§ 3)). Multiplication of $K_1$ of ($5_1$) and $C_{101}$ of § 1 by powers of $Q$ gives covariants of any odd order led by $a_1$, $a_0 + K$, and, by subtracting such covariants, covariants led by $a_0 + a_2 + a_3$.

Lemma. *There exists no covariant of odd order led by $K$.*

Assume such a covariant in the form $C = Kx_1^{2h+1} + Ix_1^{2h} x_2 + \cdots$. Then from (5), $m = 4$,

$$K(x_1^{2h+1} + x_1^{2h} x_2 t + \cdots) + I'(x_1^{2h} x_2 + \cdots) \equiv C \pmod 2,$$

where $I$ becomes $I'$ under (5). Hence, writing $I' \equiv I + tI_1$,

$$K \equiv I_1 \pmod 2.$$

Now if $I$ contains $a_4$, the increment $I_1$ contains $a_0$, whereas $K$ does not. Hence $I$ does not contain $a_4$. But the increment to a linear function of $a_0$, $a_1$, $a_2$, $a_3$ contains $a_1$ only, whereas $K$ contains $a_2$. Hence $K \not\equiv I_1$, a contradiction.

It now follows that the general form of a covariant of odd order is

$$C = [\lambda a_1 + \mu(a_0 + a_2 + a_3)]\, x_1^{2k+1} + \cdots.$$

Thus

$$C \equiv (\lambda + \mu) Q^k C_{101} + \mu Q^k K_1 + L\Gamma \pmod 2,$$

where $\Gamma$ is of even order $2k - 2$.

Any covariant of even order may be written

$$C_1 = [\lambda a_0 + \mu a_1 + \nu(a_2 + a_3)]\, x_1^{2h} + \cdots,$$

and we have, using $K_2$ ($m = 4$) of § 3, and $C_{012}$ of § 1,

$$C_1 \equiv Q^{h-1}[\lambda K_2 + (\mu + \nu) C_{012}] + \nu K Q^h + L\Gamma_1 \pmod 2;$$

$\Gamma_1$ being a first degree covariant of odd order $2h - 3$. Hence the fundamental system sought is

(10) $$K,\ K_1,\ K_2,\ C_{101},\ C_{012},\ L,\ Q.$$

The form $f_4$ itself is reducible as follows [cf. (8)]:

$$f_4 \equiv QK_2 + L(K_1 + C_{101}) \pmod 2.$$

Employing $K_1$, $K_2$, $K$, $K_{m3}$ $(m = 5)$ of § 3 and $C_{021}$, $C_{102}$ of § 1, we find for the required fundamental system of the quintic,

$$K,\ K_1,\ K_2,\ K_{53},\ C_{021},\ C_{102},\ L,\ Q. \tag{11}$$

The reduction of the form $f_5$ itself is given by

$$f_5 \equiv QK_{53} + L(K_2 + C_{102}) \pmod 2 \quad \text{[cf. (9)]}.$$

## 6. On the complete system of the cubic, modulo 2

The development of the asyzygetic theory of the concomitants of a form often precedes the derivation of its complete system. In this section we show that every covariant of order $> 3$ of a binary cubic, modulo 2, is quasi-reducible (i. e., reducible on multiplication by $K^s$ $(s \geqq 0)$) in terms of a set of fourteen concomitants, nine covariants and five invariants.

It is known that the fundamental system of seminvariants* of $f_3$ modulo 2 is

$$\begin{gathered} a_0, \quad K = a_1 + a_2, \quad \delta_{00} = (a_0 + K + a_3)\,a_3, \\ \Delta = a_0\,a_3 + a_1\,a_2, \quad \beta = a_1^2 + a_0\,a_1, \end{gathered} \tag{12}$$

and also that $f_3$ has the invariants

$$\begin{gathered} K, \quad \Delta, \quad I = a_0^2 + a_0\,K + \delta_{00}, \quad k = a_0\,\delta_{00}, \\ V = \beta(\beta + K^2 + a_0\,K)(\Delta + \delta_{00}). \end{gathered} \tag{13}$$

In two previous papers (quoted above) I have shown that $f_3$ has the covariants†

$$\begin{aligned} H &= sx_1^2 + \Delta x_1\,x_2 + (a_1\,a_3 + a_2^2)\,x_2^2, \\ G_1 &= (\Delta + s)\,x_1 + (\Delta + a_1\,a_3 + a_2^2)\,x_2 \qquad (s = a_0\,a_2 + a_1^2), \\ P &= a_0^2\,x_1^3 + a_1^2\,x_1^2\,x_2 + a_2^2\,x_1\,x_2^2 + a_3^2\,x_2^3, \\ K_1 &= (a_0 + K)\,x_1 + (K + a_3)\,x_2, \\ K_2 &= a_0\,x_1^2 + Kx_1\,x_2 + a_3\,x_2^2, \\ C_1 &= (a_0^2 + K^2)\,x_1 + (K^2 + a_3^2)\,x_2, \\ C_2 &= a_0^2\,x_1^2 + K^2\,x_1\,x_2 + a_3^2\,x_2^2. \end{aligned} \tag{14}$$

There is, therefore, a covariant led by $\beta$, viz.,

$$t = KK_2 + H = \beta x_1^2 + (\Delta + K^2)\,x_1\,x_2 + (a_2\,a_3 + a_2^2)\,x_2^2.$$

* Dickson, *Madison Colloquium Lectures*, p. 53.

† Cf. American Journal of Mathematics, loc. cit., p. 78. The forms $C_4^{(1)}$, $C_2^{(3)}$ in Table I are reducible as follows:

$$C_4^{(1)} \equiv LK_1 + QK_2, \qquad C_2^{(3)} \equiv Kt + (\Delta + K^2)\,K_2 \pmod 2.$$

None of the invariants $\Delta$, $g$ (below), $I$ can be leading coefficients of covariants of odd order, but there exists a cubic covariant led by the invariant $K$, viz.,

$$G = QK_1 + f_3 = Kx_1^3 + (a_0 + a_1 + a_3)x_1^2 x_2 + (a_0 + a_2 + a_3)x_1 x_2^2 + Kx_2^3.$$

Taking the elementary symmetric function of second degree in the coefficients of $t$, as explained in § 3 (4), we have the fourth degree invariant of $f_3$,

$$(15)\quad g = \beta^2 + \beta(\Delta + K^2) + \beta\beta_1 + \beta_1(\Delta + K^2) + \beta_1^2 \qquad (\beta_1 = a_2 a_3 + a_3^2),$$
$$\equiv \beta^2 + \beta(\Delta + K^2) + (\Delta + \delta_{00})(\beta + a_0 K + K^2) \qquad (\text{mod } 2).$$

Now from (13), (15) we have

$$(16)\quad \begin{gathered} a_0^3 + a_0^2 K + a_0 I + k \equiv 0, \\ \beta^2 + \beta(a_0^2 + a_0 K + I + K^2) \\ + (a_0^2 + a_0 K + I + \Delta)(a_0 K + K^2) + g \equiv 0. \end{gathered}$$

Any seminvariant $\phi$ of $f_3$, being a polynomial in the seminvariants (12), is a polynomial in $a_0, K, \beta, \Delta, I$;

$$\phi = \phi(a_0, \beta, K, \Delta, I).$$

We now use congruences (16) as reducing moduli, whereupon we are able to reduce all exponents* of $\beta$ (in $\phi$) below 2 and all exponents of $a_0$ below 3. Thus any seminvariant can be reduced to the form

$$\phi = J_0 + J_1 a_0 + J_2 a_0^2 + (\Gamma_0 + \Gamma_1 a_0 + \Gamma_2 a_0^2)\beta,$$

wherein $J_i$, $\Gamma_i$ $(i = 0, 1, 2)$ are invariants (expressed in terms of $K, \Delta, k, I, g$) some of which might be zero. Let $C_M$ be any covariant of $f_3$ of even order $M = 2h$, led by $\phi$; $C_M = \phi x_1^{2h} + \cdots$. Then,

$$C_M \equiv Q^h J_0 + Q^{h-1}(K_2 J_1 + C_2 J_2) + Q^{h-1} t\Gamma_0 + Q^{h-2}(K_2 \Gamma_1 + C_2 \Gamma_2)t + LC \qquad (\text{mod } 2),$$

where $C$ is a covariant of odd order $2h - 3$. Thus every covariant of even order $> 3$ is reducible in terms of our covariants and invariants.

Proceeding to covariants of odd order; there is a covariant with seminvariant leader $B = \beta + \Delta$, viz.,

$$l = Bx_1^3 + (a_0 a_2 + a_1 a_2 + a_1 a_3 + a_2^2)x_1^2 x_2 + (a_0 a_2 + a_1 a_2 + a_1 a_3 + a_1^2)x_1 x_2^2 + (\Delta + a_2 a_3 + a_2^2)x_2^3 \equiv QG_1 + Kf_3.$$

* Cf. L. J. Reed, *"Some fundamental systems of formal modular invariants and covariants,"* Dissertation, University of Pennsylvania, 1915, § 3.

Replacing $\beta$ by $B + \Delta$ in (16) and in $\phi$ we can reduce any seminvariant to the form

$$\phi = R_0 + R_1 a_0 + R_2 a_0^2 + (S_0 + S_1 a_0 + S_2 a_0^2) B,$$

where $R_i$, $S_i$ $(i = 0, 1, 2)$ are polynomials in the five invariants $K$, $\Delta$, $k$, $I$, $g$. Let $C_N$ be a covariant of $f_3$ of odd order $N = 2h + 3 > 3$, led by $\phi$; $C_N = \phi x_1^N + \cdots$. Then

$$KC_N \equiv Q^h R_0 G + KR_1 Q^h f_3 + KR_2 Q^h P$$
$$+ K(lS_0 Q^h + lS_1 K_2 Q^{h-1} + lS_2 C_2 Q^{h-1}) + LC' \pmod 2.$$

We have now proved the following:*

THEOREM: *Every formal covariant modulo 2 of order $> 3$ of the binary cubic $f_3$ is quasi-reducible, upon multiplication by $K^s$ $(s \geqq 0)$, in terms of the fourteen concomitants listed below:*

$$K, \Delta, k, I, g, f_3, K_2, C_2, K_1, G_1, P, t, L, Q.$$

*On the subject of quasi-reducibility in ternariant theory, cf. Forsyth, American Journal of Mathematics, vol. 12 (1890).

UNIVERSITY OF PENNSYLVANIA